

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 2, 2004, 18:28:29 ; Search time 2829 Seconds
(without alignments)
3998.774 Million cell updates/sec

Title: US-10-063-731-118
Perfect score: 1357
Sequence: 1 MSTTTCQVAVFLLSILGLAG.....DGGARTEDEVQSPSKHDYV 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_i/uspto spool/US10063731/runat_01092004_155040_18714/app query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10063731@cgn 1 1 2527 @runat_01092004_155040_18714 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *

29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rod: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|--------|---------------|--------|----|-----------|--------------------|
| 1 | 1357 | 100.0 | 816 | 9 | AF221069 | AF221069 Homo sapi |
| 2 | 1357 | 100.0 | 982 | 6 | BD237992 | BD237992 Gastric p |
| 3 | 1357 | 100.0 | 1869 | 6 | AX472954 | AX472954 Sequence |
| 4 | 1357 | 100.0 | 2121 | 6 | AX092386 | AX092386 Sequence |
| 5 | 1357 | 100.0 | 2121 | 6 | AX376288 | AX376288 Sequence |
| 6 | 1357 | 100.0 | 2121 | 6 | AX697257 | AX697257 Sequence |
| 7 | 1357 | 100.0 | 2121 | 9 | AY358479 | AY358479 Homo sapi |
| 8 | 1264 | 93.1 | 796 | 9 | AF349452 | AF349452 Homo sapi |
| 9 | 1264 | 93.1 | 3344 | 9 | AK098474 | AK098474 Homo sapi |
| 10 | 1235.5 | 91.0 | 2786 | 10 | AF221068 | AF221068 Mus muscu |
| 11 | 1217.5 | 89.7 | 1409 | 10 | AF349450 | AF349450 Mus muscu |
| 12 | 1165.5 | 85.9 | 849 | 10 | AF349451 | AF349451 Mus muscu |
| 13 | 1147.5 | 84.6 | 861 | 10 | AF349453 | AF349453 Mus muscu |
| 14 | 1120 | 82.5 | 783 | 6 | BD237993 | BD237993 Gastric p |
| 15 | 1107.5 | 81.6 | 2040 | 9 | BC063845 | BC063845 Homo sapi |
| 16 | 1011.5 | 74.5 | 2388 | 5 | BC060425 | BC060425 Xenopus l |
| 17 | 397 | 29.3 | 164953 | 9 | AC016252 | AC016252 Homo sapi |
| 18 | 394 | 29.0 | 423 | 6 | AX380609 | AX380609 Sequence |
| 19 | 394 | 29.0 | 423 | 6 | AX380725 | AX380725 Sequence |
| 20 | 389 | 28.7 | 145477 | 2 | AC010810 | AC010810 Homo sapi |
| 21 | 388 | 28.6 | 636 | 6 | E31608 | E31608 Tight junct |
| 22 | 388 | 28.6 | 636 | 10 | AF087825 | AF087825 Mus muscu |
| 23 | 388 | 28.6 | 894 | 10 | BC008104 | BC008104 Mus muscu |
| 24 | 388 | 28.6 | 1235 | 10 | BC050007 | BC050007 Mus muscu |
| 25 | 388 | 28.6 | 158198 | 2 | AC012013 | AC012013 Homo sapi |
| 26 | 385 | 28.4 | 888 | 10 | AF486651 | AF486651 Mus muscu |
| 27 | 382 | 28.2 | 2859 | 6 | AX833932 | AX833932 Sequence |
| 28 | 382 | 28.2 | 2859 | 9 | AK096063 | AK096063 Homo sapi |
| 29 | 375.5 | 27.7 | 690 | 6 | AX921675 | AX921675 Sequence |
| 30 | 375.5 | 27.7 | 1616 | 9 | BC030524 | BC030524 Homo sapi |
| 31 | 373 | 27.5 | 245225 | 2 | AC098329 | AC098329 Rattus no |
| 32 | 373 | 27.5 | 247885 | 2 | AC097923 | AC097923 Rattus no |
| 33 | 370 | 27.3 | 208814 | 2 | AC134606 | AC134606 Mus muscu |
| 34 | 370 | 27.3 | 261333 | 2 | AC138280 | AC138280 Mus muscu |
| 35 | 368.5 | 27.2 | 1207 | 9 | HSA011497 | AJ011497 Homo sapi |
| 36 | 368.5 | 27.2 | 1253 | 6 | AR400627 | AR400627 Sequence |
| 37 | 368.5 | 27.2 | 1253 | 6 | AX013131 | AX013131 Sequence |
| 38 | 368.5 | 27.2 | 1253 | 6 | BD206312 | BD206312 Human nuc |
| 39 | 368.5 | 27.2 | 1546 | 6 | BD234896 | BD234896 Human apo |
| 40 | 368.5 | 27.2 | 1711 | 6 | BD063242 | BD063242 Secreted |
| 41 | 368.5 | 27.2 | 1787 | 6 | AX098214 | AX098214 Sequence |
| 42 | 368 | 27.1 | 681 | 6 | AX276968 | AX276968 Sequence |
| 43 | 367 | 27.0 | 580 | 10 | AF249889 | AF249889 Mus muscu |
| 44 | 364.5 | 26.9 | 678 | 6 | AX083410 | AX083410 Sequence |
| 45 | 364.5 | 26.9 | 1095 | 6 | AX083420 | AX083420 Sequence |

ALIGNMENTS

Alignment Scores: 3.49e-126 Length: 982
Pred. No.: 1357.00 Matches: 261
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

US-10-063-731-118 (1-261) x BD237992 (1-982)

QY 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 70 ATGTCCACCACACATGCCAAGTGTGGCGTTCTCTCTGTCATCTGGGGCTGGCCGC 129
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 130 TGCATCGCGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACAAACCCGTC 189
QY 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
Db 190 ACCTCCGTGTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTTCAGGCTTC 249
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 250 ACCGAATGAGGCCCTATTTCACCATCTGGGACTTCCAGCCATGCTGCAGGCAGTCGA 309
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 310 GCCCTGATGATCGTAGGCATCGTCTGGTGCCATTGGCCTCTGGTATCCATCTTTGCC 369
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 370 CTGAATGCATCGCATTTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACACTGACC 429
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 430 TCCGGGATCATGTTTCATGTTCTCAGGTCCTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 489
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 490 AACATGCTGGTGACTAACTTCTGGATGTCACAGCTAACATGTACACCGGCATGGGTGG 549
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 550 ATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGCGGCTCTGTTCGGGCTGGGTC 609
QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 610 GCTGGAGGCCTCACACTAAATTGGGGGTGTGATGATGTGATCGCTCGCCGGGGCTGGCA 669
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 670 CCAGAAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTTCCTAC 729
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
Db 730 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCACACCAAAAAACAAGAATA 789
QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 790 TAGCATGGAGGTGCCCCGACAGAGGACGAGGTACAATCTTATCTTCTTCCAGCACGACTAT 849
QY 261 Val 261
Db 850 GTG 852

RESULT 3
AX472954
LOCUS AX472954 1869 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 3 from Patent WO0218576.
ACCESSION AX472954
VERSION AX472954.1 GI:22207741
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chen,S.Y., Macina,R.A., Sun,Y. and Recipon,H.
TITLE Compositions and methods relating to lung specific genes
JOURNAL Patent: WO 0218576-A 3 07-MAR-2002;
Diadexus, Inc. (US)
FEATURES
Location/Qualifiers
source 1..1869
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores: 7.96e-126 Length: 1869
Pred. No.: 1357.00 Matches: 261
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-10-063-731-118 (1-261) x AX472954 (1-1869)

QY 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 83 ATGTCCACCACACATGCCAAGTGTGGGTTCCTCTCTGTCATCTGGGGCTGGCCGC 142
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 143 TGCATCGCGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACAAACCCGTC 202
QY 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
Db 203 ACCTCCGTGTTCAGTACGAGGGCTCTGGAGGAGCTCGGTGAGGCAGATTTCAGGCTTC 262
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 263 ACCGAATGCAGGCCCTATTTCACCATCTGGGACTTCCAGCCATGCTGCAGGCAGTCGA 322
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 323 GCCCTGATGATCGTAGGCATCGTCTGGTGCCATTGGCCCTCTGGTATCCATCTTTGCC 382
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 383 CTGAATGCATCCGCATTTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACACTGACC 442
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 443 TCCGGGATCATGTTTCATGTTCTCAGGTCCTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 502
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 503 AACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG 562
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 563 ATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGGCTCTGTTCGTGGGCTGGGTC 622
QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 623 GCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGATCGCTCGCCGGGCTGGCA 682
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 683 CCAGAAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTTCCTAC 742
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
Db 743 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAAACAAGAGATA 802

Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 803 TAGGATGGAGGTGCCCCGACAGAGGACGAGGTACAAATCTTATCTTCCCAAGCAGACTAT 862

Qy 261 Val 261
Db 863 GTG 865

RESULT 4
AX092386
LOCUS AX092386 2121 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 117 from Patent WO0116318.
ACCESSION AX092386
VERSION AX092386.1 GI:13444507
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A.,
Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and
Wood,W.I.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0116318-A 117 08-MAR-2001;
Genentech, Inc. (US)
FEATURES
source Location/Qualifiers
1. .2121
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 9.37e-126 Length: 2121
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-063-731-118 (1-261) x AX092386 (1-2121)

Qy 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 90 ATGTCCACCACCACATGCCAAGTGGTGGCTCCTGTCCATCCTGGGCTGGCCGGC 149

Qy 21 CysIleAlaAlaThrGlyMetAspMetTyrProSerThrGlnAspLeuTyrAspAsnProVal 40
Db 150 TGCATCGCGGCCACCGGGATGGACATGTGGAGACCCAGGACCTGTACGACAAACCCCGTC 209

Qy 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
Db 210 ACCTCCGTGTTCAGTAGCAAGGGCTCTGGAGGAGCTGGTGAGCAGAGTTCAGGCTTC 269

Qy 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 270 ACCGAATGCAGGCCCTATTTCACCATCTGGACTTCCAGCCATGCTGCAGGCAGTGCGA 329

Qy 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 330 GCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTGGCTCCTGGTATCCATCTTTGCC 389

Qy 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 390 CTGAATGCATCCGCAATGGCAGCATGGAGACTCTGCCAAGCAACATGACACTGACC 449

Qy 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 450 TCCGGGATCATGTTTCATTTGTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 509

Qy 141 AenMetLeuValThrAsnPheTyrMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 510 AACATGCTGGTGACTAACTTCTGGATGTCCACAGTAAACATGTACACCGCATGGGTGG 569

Qy 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTyrVal 180
Db 570 ATGGTGCAAGTGTTCAGACCAGGTACACATTTGGTGGCTCTGTTCGTGGCTGGGTC 629

Qy 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 630 GCTGGAGGCTCACACTAATTGGGGTGTGTATGATGTGCATGCCTGCCGGGCGCTGGCA 689

Qy 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 690 CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTGCCTAC 749

Qy 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
Db 750 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAACAAGAAGATA 809

Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 810 TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAAATCTTATCTTCAAGCAGACTAT 869

Qy 261 Val 261
Db 870 GTG 872

RESULT 5
AX376288
LOCUS AX376288 2121 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 355 from Patent WO0168848.
ACCESSION AX376288
VERSION AX376288.1 GI:19170535
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Baker,K.P., Chen,J., Desnoyers,L., Goddard,A., Godowski,P.J.,
Gurney,A.L., Pan,J., Smith,V., Watanabe,C.K., Wood,W.I. and
Zhang,Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0168848-A 355 20-SEP-2001;
Genentech, Inc. (US)
FEATURES
source Location/Qualifiers
1. .2121
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 9.37e-126 Length: 2121
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-063-731-118 (1-261) x AX376288 (1-2121)

Qy 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 90 ATGTCCACCACCACATGCCAAGTGGTGGCTTCTCTCTGTCCATCCTGGGCTGGCCGGC 149

Qy 21 CysIleAlaAlaThrGlyMetAspMetTyrProSerThrGlnAspLeuTyrAspAsnProVal 40
Db 150 TGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAAACCCCGTC 209

Qy 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60

Db 210 ACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTC 269
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 270 ACCGAATGCAGGCCCTATTTCACCATCTGGGACTTCCAGCCATGCTGCAGGCAGTGC GA 329
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuValSerIlePheAla 100
Db 330 GCCCTGATGATCGTAGGCATCGTCTGGTGCCATTGGCCTCCTGGTATCCATCTTTGCC 389
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 390 CTGAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACACTGACC 449
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 450 TCCGGGATCATGTTTCATTGTTCTCAGGTTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 509
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 510 AACATGCTGGTGACTAACTTCTGGATGTCACAGCTAACATGTACACCGGCATGGGTGGG 569
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 570 ATGGTGCAGACTGTTCCAGACCAGGTACACATTTGGTGGCTCTGTTCTGGTGGCTGGGTC 629
QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 630 GCTGGAGGCCCTCACACTAATTGGGGTGTGATGATGTGCATCGCCTGCCGGGCTGGCA 689
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 690 CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTTCGCTAC 749
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
Db 750 AAGCCTGGAGGCTTCAAGCCAGCACTGGCTTTGGTCCAAACACCAAAAACAAGAAGATA 809
QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 810 TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAAATCTTATCCTTCCAAGCACGACTAT 869
QY 261 Val 261
Db 870 GTG 872

RESULT 6
AX697257
LOCUS
DEFINITION Sequence 325 from Patent WO0078961.
ACCESSION AX697257
VERSION AX697257.1 GI:29498417
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ferrara,N., Stewart,T.A., Williams,P.M., Baker,K.P., Desnoyers,L.,
Eaton,D.L., Gao,W.Q., Pan,J., Botstein,D., Fong,S., Goddard,A.,
Godowski,P.J., Gurney,A.L., Smith,V., Tumas,D., Wood,W.I.,
Grimaldi,C.J., Hillan,K.J., Paoni,N.F., Roy,M.A. and Watanabe,C.K.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0078961-A 325 28-DEC-2000;
Genentech Inc. (US)
FEATURES
source Location/Qualifiers
1. .2121
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores: 9.37e-126 Length: 2121
pred. No.: 1357.00 Matches: 261
Score: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-063-731-118 (1-261) x AX697257 (1-2121)
QY 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 90 ATGTCCACCACACATGCCAAGTGGTGGCGTTCTCTGTCTCCATCCTTGGGGCTGGCCGC 149
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 150 TGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTC 209
QY 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
Db 210 ACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTCGTGAGGCAGAGTTCAGGCTTC 269
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 270 ACCGAATGCAGGCCCTATTTCACCATCTCTGGGACTTCCAGCCATGCTGCAGGCAGTGC GA 329
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuValSerIlePheAla 100
Db 330 GCCCTGATGATCGTAGGCATCGTCTGGTGCCATTGGCCTCCTGGTATCCATCTTTGCC 389
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 390 CTGAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACACTGACC 449
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 450 TCCGGGATCATGTTTCATTGTTCTCAGGTTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 509
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 510 AACATGCTGGTGACTAACTTCTGGATGTCACAGCTAACATGTACACCGGCATGGGTGGG 569
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 570 ATGGTGCAGACTGTTCCAGACCAGGTACACATTTGGTGGCTCTGTTCGTGGGCTGGGTC 629
QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 630 GCTGGAGGCCCTCACACTAATTGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCA 689
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 690 CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTTCGCTAC 749
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
Db 750 AAGCCTGGAGGCTTCAAGCCAGCACTGGCTTTGGTCCAAACACCAAAAACAAGAAGATA 809
QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 810 TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAAATCTTATCCTTCCAAGCACGACTAT 869
QY 261 Val 261
Db 870 GTG 872
RESULT 7
AX358479
LOCUS
DEFINITION Homo sapiens clone DNA73734 CLDN18 (UNQ778) mRNA, complete cds.
ACCESSION AX358479
VERSION AX358479.1 GI:37182080

KEYWORDS FLI_CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2121)
Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,
Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B.,
Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E.,
Heidens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S.,
Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C.,
Schoenfeld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,V.,
Stinson,J., Vagts,A., Vandlen,R., Watanabe,C., Wieand,D., Woods,K.,
Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z.,
Goddard,A., Wood,W.I. and Godowski,P.
TITLE The Secreted Protein Discovery Initiative (SPDI), a Large-Scale
Effort to Identify Novel Human Secreted and Transmembrane Proteins:
A Bioinformatics Assessment
JOURNAL Genome Res. 13 (10), 2265-2270 (2003)
PUBMED 12975309
REFERENCE 2 (bases 1 to 2121)
AUTHORS Clark,H.F.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,
Inc., 1 DNA Way, South San Francisco, CA 94080, USA
FEATURES
source
1..2121
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DNA73734"
1..2121
/locus_tag="UNQ778"
90..875
/locus_tag="UNQ778"
/note="PRO1572"
/codon_start=1
/product="CLDN18"
/protein_id="AAQ88843.1"
/db_xref="GI:37182081"
/translation="MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPTVSVP
QYGLWRSCVRQSSGFTECPYFTILGLPALQAVRALMIVGIVLGAIGLLVSTFALK
CIRIGMEDSAKANWLTLSGIMFIVSGLCAIAGVSFANMLVTFNFMSTANMYTGMGG
MVQTVQTRYTFGAALFVGWVAGGLTLIGVMMCIACRLAPEETNYKAVSYHASGHSV
AYKPGGFKASTGFGSNTKXKIYDGGARTEDEVQSPSKHDYV"
ORIGIN
Alignment Scores:
Pred. No.: 9.37e-126 Length: 2121
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-063-731-118 (1-261) x AY358479 (1-2121)
QY 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 90 ATGTCCACCACCATGCCAAGTGGTGGCTTCTCTCTCCATCTGGGGCTGGCCGGC 149
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 150 TGCATCGCGGCCACCGGATGGACATGTGAGCACCAGGACCTGTACGACCAACCCCGTC 209
QY 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
Db 210 ACCTCCGTGTTCAGTACGAAGGGCTCTGAGGAGCTGCGTGAGGCAGAGTTTCAGGCTTC 269
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 270 ACCGAATGCAGGCCCTATTTCACCATCTGGGACTTCCAGCCATCTGCAGGCAGTGCGA 329

QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 330 GCCCTGATGATCGTAGGCATCGTCTCTGGTGCCATGGCCCTCTGGTATCCATCTTTGCC 389
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 390 CTGAAATGCATCCGCAATTGGCAGCATGGAGGACTCTGCCAAAGCCACATGACACTGACC 449
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 450 TCCGGGATCATGTTTCATGTCTCTCAGGTCTTTGTGCAATTTGCTGGAGTGTGTGTTTGGC 509
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 510 AACATGCTGGTGACTAACTTCTGGATGTCCAGCTAACATGTACACCGCATGGGTGG 569
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 570 ATGGTGCACTGTTTCAGACCCAGGTACACATTTGGTGGGCTCTGTTGGTGGGCTGGGTC 629
QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 630 GCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCA 689
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 690 CCAGAAAGAAACCACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTAC 749
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
Db 750 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTGGGTCCAAACACCAAAAAACAAGAATA 809
QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 810 TACGATGGAGTGTCCCGCACAGAGGACGAGTACATCTTATCTCTTCCAAGCAGACTAT 869
QY 261 Val 261
Db 870 GTG 872
RESULT 8
AF349452
LOCUS Homo sapiens claudin-18A2.1 mRNA, complete cds, alternatively
DEFINITION spliced.
ACCESSION AF349452
VERSION AF349452.1 GI:16224168
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 796)
Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B.,
Popescu,N.C. and Kimura,S.
TITLE claudin-18, a novel downstream target gene for the T/EBP/NKX2.1
homeodomain transcription factor, encodes lung- and
stomach-specific isoforms through alternative splicing
JOURNAL Mol. Cell. Biol. 21 (21), 7380-7390 (2001)
MEDLINE 21470315
PUBMED 11585919
REFERENCE 2 (bases 1 to 796)
AUTHORS Niimi,T. and Kimura,S.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2001) Laboratory of Metabolism, National Cancer
Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
1..796
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
10..795
/note="alternatively spliced"
CDS


```
Qy 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 234 ACCGAGTCCGGGGCTACTTCACCCCTGCTGGGGCTGCCAGCCATGCTGCAGGCAGTGGCA 293
Qy 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 294 GCCTGATGATCGTAGGCATGCTCCTGGTGCCATGGCCTCCTGGTATCCATCTTTGCC 353
Qy 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 354 CTGAATGCATCCGCATTCGCAGCATGGAGGACTCTGCCAAAGCCAACATGACACTGACC 413
Qy 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 414 TCCGGGATCATGTTTCATGCTCAGGTCTTTGTGCAATTGCTGGAGTGCTGTGTTGCC 473
Qy 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 474 AACATGCTGGTGACTAACTTCTGGATGTCCACAGTAACATGTACACCGCATGGGTGGG 533
Qy 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 534 ATGGTGACAGACTGTTCCAGACCAGGTACACATTTGGTGGGCTCTGTTCTGGGCTGGGTC 593
Qy 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 594 GCTGGAGGCCTCACACTAATTGGGGTGTGATGTGCATCGCTCGCGGGCCCTGGCA 653
Qy 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 654 CCAGAAGAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTTCCTAC 713
Qy 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
Db 714 AAGCCTGGAGGCTTCAAGCCAGCAGCTGGCTTTGGGTCCAACACCAAAACAAGAAGATA 773
Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 774 TACGATGAGGTGCCCGCACAGAGGACGAGGTACAAATCTTATCCTTCCAAGCAGCATAT 833
Qy 261 Val 261
Db 834 GTG 836
RESULT 10
AF221068
LOCUS Mus musculus Claudin-18 mRNA, complete cds.
DEFINITION AF221068
ACCESSION AF221068
VERSION AF221068.1 GI:6715515
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2786)
AUTHORS Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B.,
Popescu,N.C. and Kimura,S.
TITLE claudin-18, a novel downstream target gene for the T/EBP/NKX2.1
homeodomain transcription factor, encodes lung- and
stomach-specific isoforms through alternative splicing
JOURNAL Mol. Cell. Biol. 21 (21), 7380-7390 (2001)
MEDLINE 21470315
PUBMED 11585919
REFERENCE 2 (bases 1 to 2786)
AUTHORS Niimi,T. and Kimura,S.
TITLE Direct Submission
JOURNAL Submitted (04-JAN-2000) Laboratory of Metabolism, DBS, National
Cancer Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
1. .2786
/organism="Mus musculus"
```

CDS

```
/mol_type="mRNA"
/db_xref="taxon:10090"
68. .862
/codon_start=1
/product="Claudin-18"
/protein_id="AAF26447.1"
/db_xref="GI:6715516"
/translation="MATTCQVVGVLLSLGLAGCIAATGMDMWSTQDLYDNPVAVF
QYEGLRSCVQSSSGFTECRPYFTILGLPAMLQAVRALMIVGIVLGVILVSI FALK
CIRIGSMDDSAKAKMTLTSGILFTISGICAIIGSVFANMLVTNFMWSTANMYSGMGG
MGVMVQTVQTRYTFGAALFVGWVAGGLTLIGVMMCIAACRGLTPDDSNFKAVSYHAGS
QNVAYRPGGPKASTGFGSNTRNKIYDGGARTEDDEQSHPTKYDYV"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 1.84e-113 Length: 2786
Score: 1235.50 Matches: 232
Percent Similarity: 95.83% Conservative: 21
Best Local Similarity: 87.88% Mismatches: 8
Query Match: 91.05% Indels: 3
DB: 10 Gaps: 1
US-10-063-731-118 (1-261) x AF221068 (1-2786)
Qy 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 68 ATGGCCACCACCACGTCGCCAGGTGGTAGGGCTTCTCCTGTCCCTCCTGGGTCTGGCCGC 127
Qy 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 128 TGCATAGCCGCCACTGGGATGGACATGTGGAGCACTCAAGACCTGTATGACACCCAGTC 187
Qy 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
Db 188 ACCGCCGTGTTCCAGTATGAAGGCTCTGGAGGAGTTGCGTGCAACAGAGCTCGGGTTC 247
Qy 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 248 ACCGAGTGCCGGCCATACTTCAACATCCTGGGCCCTTCCAGCCATGCTGCAAGCTGTACGA 307
Qy 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 308 GCCCTGATGATCGTGGGCATTGTTCTGGGGGTTCATCGGTATCCTCGTGTCCATCTTCGCC 367
Qy 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 368 CTGAAGTGCAATTCGCATTGGTAGCATGGATGACTCTGCCAAGSCCAAGATGACTCTGACT 427
Qy 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 428 TCTGGGATCTTGTTCATCATCTCCGGCATCTGTGCAATCATTTGGTGTGTCTGTGTTGCC 487
Qy 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThr-----Gly 157
Db 488 AACATGCTGGTGACCAACTTCTGGATGTCCACAGCTAACATGTACAGCGGCATGGCGGC 547
Qy 158 MetGlyGlyMetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheVal 177
Db 548 ATGGGTGGCATGTGTCAGACCCGTTTCAGACCAGGTACACCTTCGGTGCAGCTCTGTTCGTG 607
Qy 178 GlyTrpValAlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArg 197
Db 608 GGCTGGGTGCTGGAGGCCTCACCCCTGATTGGGGAGTGTATGATGATGATGATGATGATGATG 667
Qy 198 GlyLeuAlaProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSer 217
Db 668 GGCCTGACACCAGATGACAGCAACTTCAAAGCTGTGTCTTACCATGCTCTTGGCCAAAT 727
Qy 218 ValAlaTyrLysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsn 237
Db 728 GTTGCTACAGGCCTGGAGGCTTTAAGGCCAGCACTGGCTTTGGTTCACACACCAAGAAC 787
Qy 238 LysLysIleTyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLys 257
```


||||| 788 AAGAAGATCTACGATGGGGTGCCTCCGACAGAACGATGAACAGTCTCATCTACCAAG 847
QY 258 HisAspTyrVal 261
:::|||||
Db 848 TATGACTATGTG 859

RESULT 11
AF349450 1409 bp mRNA linear ROD 18-OCT-2001
LOCUS Mus musculus claudin-18A1.2 mRNA, complete cds, alternatively
DEFINITION spliced.
ACCESSION AF349450
VERSION AF349450.1 GI:162224162
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1409)
AUTHORS Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B.,
Popescu,N.C. and Kimura,S.
TITLE homeodomain transcription factor, encodes lung- and
stomach-specific isoforms through alternative splicing
JOURNAL Mol. Cell. Biol. 21 (21), 7380-7390 (2001)
MEDLINE 21470315
PUBMED 11585919
REFERENCE 2 (bases 1 to 1409)
AUTHORS Niimi,T. and Kimura,S.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2001) Laboratory of Metabolism, National Cancer
Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
1. .1409
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/note="isolated from mixed background of 129/Sv, C57BL/6,
and Black Swiss mice"
62. .688
/note="alternatively spliced"
/codon_start=1
/product="claudin-18A1.2"
/protein_id="AAL15635.1"
/db_xref="GI:162224163"
/translation="MATTCQVVGLLLSLLGLAGCIAATGMDMWSTQDLYDNPVTAVF
QYEGLRSCVQSSGFTECRPYFTILGLPAMLOAVRALMIVGIVLGIVIGILVSIFALK
CIRIGSMDDSAKAMTILTSGLIFISGICAIIGVSVFANMLVTNFMWSTANNYSYSGMG
MGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIACRGLTPDDSK"

ORIGIN
Alignment Scores:
Pred. No.: 4.83e-112 Length: 1409
Score: 1217.50 Matches: 231
Percent Similarity: 94.03% Conservative: 21
Best Local Similarity: 86.19% Mismatches: 9
Query Match: 89.72% Indels: 7
DB: 10 Gaps: 2

US-10-063-731-118 (1-261) x AF349450 (1-1409)

QY 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
|||||
Db 62 ATGGCCACCACCACCGTCCAGGTGGTAGGGCTTCCTCTCCCTCTGGGTCTGGCCGGC 121
|||||

QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAsnProVal 40
|||||

Db 122 TGCATAGCCGCCACTGGGATGGACATGTGGAGCACTCAAGACCTGTATGACAAACCCAGTC 181
|||||

QY 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
|||||

Db 182 ACCGCCGTGTTCCAGTATGAAGGGCTCTGGAGGAGTTGCGTGCAACAGAGCTCGGGGTTTC 241
|||||

QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
|||||

Db 242 ACCGAGTGCCGGCCATACCTTCACCATCTCTGGGCGCTTCCAGCCATGCTGCAAGCTGTACGA 301
|||||

QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
|||||

Db 302 GCCCTGATGATCGTGGGCATTGTTCTGGGGGTTCATCGGTATCTCTCGTGTCCATCTTCGCC 361
|||||

QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
|||||

Db 362 CTGAAGTGCAATTGTCATTTGGTAGCATGGATGACTCTGCCAAGSCAAGATGACTCTGACT 421
|||||

QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
|||||

Db 422 TCTGGGATCTTGTTCATCATCTCCGGCATCTGTGCAATCATTTGTTGTCTGTGTTTGCC 481
|||||

QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThr-----Gly 157
|||||

Db 482 AACATGCTGGTGACCAACTTCTGGATGTCCACAGCTAACATGATACAGCGGCATGGGCGGC 541
|||||

QY 158 MetGlyGlyMetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheVal 177
|||||

Db 542 ATGGTGGCATGTGCAGACCGTTTCAGACAGGTACACCTTCGTCAGCTCTGTTCGTG 601
|||||

QY 178 GlyTrpValAlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArg 197
|||||

Db 602 GGCTGGGTTGCTGGAGGCCTCACCCCTGATTGGGGAGTGATGATGCATCGCCTGCCGT 661
|||||

QY 198 GlyLeuAlaProGluGluThrAsn-----TyrLysAlaValSerTyrHisAla 213
|||||

Db 662 GGCCTGACACCAAGATGACAGCAAGTGAGTCTCCCTTCAAAGCTGTGTCTTACCATGCC 721
|||||

QY 214 SerGlyHisSerValAlaTyrLysProGlyGlyPheLysAlaSerThrGlyPheGlySer 233
|||||

Db 722 TCTGGCCAAATGTTGCCTACAGGCCTGGAGGCTTTAAGGCCAGCACTGGCTTTGGGTCC 781
|||||

QY 234 AsnThrLysAsnLysLysIleTyrAspGlyGlyAlaArgThrGluAspGluValGlnSer 253
|||||

Db 782 AACACCAGAAACAAGAAGATCTACGATGGGGGTGCCCGCACAGAACGATGAACAGTCT 841
|||||

QY 254 TyrProSerLysHisAspTyrVal 261
:::|||||

Db 842 CATCTACCAAGTATGACTATGTG 865
|||||

RESULT 12
AF349451 849 bp mRNA linear ROD 18-OCT-2001
LOCUS Mus musculus claudin-18A2.1 mRNA, complete cds, alternatively
DEFINITION spliced.
ACCESSION AF349451
VERSION AF349451.1 GI:162224165
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 849)
AUTHORS Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B.,
Popescu,N.C. and Kimura,S.
TITLE claudin-18, a novel downstream target gene for the T/EBP/NKX2.1
homeodomain transcription factor, encodes lung- and
stomach-specific isoforms through alternative splicing
JOURNAL Mol. Cell. Biol. 21 (21), 7380-7390 (2001)
MEDLINE 21470315
PUBMED 11585919
REFERENCE 2 (bases 1 to 849)
AUTHORS Niimi,T. and Kimura,S.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2001) Laboratory of Metabolism, National Cancer
Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
1. .849

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/note="isolated from mixed background of 129/Sv, C57BL/6, and Black Swiss mice"
54..848
/note="alternatively spliced"
/codon_start=1
/product="claudin-18A2.1"
/protein_id="AAL15636.1"
/db_xref="GI:16224166"
/translation="MSVTACQGLGFVVSLIGFAGIIAATCMDQWSTQDLYNNPVTAVF NYQGLWRSCVRESSGFTECRGYFTLLGLPAMLQAVRALMIVGLVIGILVSIPLAK CIRIGSMDDSAKAKMTLTSGILFIISGICAIIGVSVFANMLVTNFWMSTANMYSGMGG MGGMVQTVQTRYTFGAALFVGWVAGGLTLIGVMMCIACRGLTPDDSNFKAVSHASG QNVAYRPGGPKASTGFGSNTRNKIYDGGARTEDDEQSHPTKYDYV"

ORIGIN

Alignment Scores:
Pred. No.: 3.97e-107 Length: 849
Score: 1165.50 Matches: 218
Percent Similarity: 92.80% Conservative: 27
Best Local Similarity: 82.58% Mismatches: 16
Query Match: 85.89% Indels: 3
DB: 10 Gaps: 1

US-10-063-731-118 (1-261) x AF349451 (1-849)

QY 1 MetSerThrThrCysGlnValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 54 ATGTGGTGACCGCCTGCCAGGGCTTGGGGTTGTGTCACTGATCGGGTTGCGGCG 113
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 114 ATCATTCAGCCACTTGTATGGACCAGTGGAGCACCCAGGATTTATACAACACCGCGT 173
QY 41 ThrSerValPheGlnTyrGluGlyLeuTyrPArgSerCysValArgGlnSerSerGlyPhe 60
Db 174 ACCGCTGATTCAACTACCAAGGGCTATGGCGTTCATGCGTCCGAGAGAGCTCTGGCTTC 233
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 234 ACCGAGTGCCGAGGTACTTCACCCCTGTTGGGGTTGCCAGCCATGCTGCAAGCTGTACGA 293
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 294 GCCCTGATGATCGTGGGCATTGTTCTGGGGTTCATCGGTATCCTCGTCCATCTTCGCC 353
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 354 CTGAAGTGCAATCGCATTTGGTAGCATGTGACTCTGCCAAGGCCAAGATGACTCTGACT 413
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 414 TCTGGGATCTTGTTCATCATCTCCGGCATCTGTGCAATCATTTGGTGTGTCTGTTTGCC 473
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThr-----Gly 157
Db 474 AACATGCTGGTGACCAACTTCTGGATGTCCAGACTAACATGTACCGGCATGGGCGGC 533
QY 158 MetGlyGlyMetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheVal 177
Db 534 ATGGGTGGCATGGTGCAGACCGTTCCAGACCAGGTACACCTTCGGTGCAGCTCTGTTCCGTG 593
QY 178 GlyTrpValAlaGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArg 197
Db 594 GCCTGGGTTGCTGGAGCCCTCACCCCTGATTTGGGGAGTGTATGATGTGCATCGCCTGCCGT 653
QY 198 GlyLeuAlaProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSer 217
Db 654 GGCCTGACACCAGATGACAGCAACTTCAAGCTGTGCTTTACCATGCTCTTGCCAAAT 713
QY 218 ValAlaTyrLysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsn 237

Db 714 GTTGCTACAGGCTGGAGGCTTTAAGCCAGCAGCTGGCTTTGGTCCACACCAAGAAC 773
QY 238 LysLysIleTyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLys 257
Db 774 AAGAAGATCTACGATGGGGGTGCCCCGACACAGAAGACGATGAACAGTCTCATCTACCAAG 833
QY 258 HisAspTyrVal 261
Db 834 TATGACTATGTG 845

RESULT 13
AF349453

LOCUS Mus musculus claudin-18A2.2 mRNA linear ROD 18-OCT-2001
DEFINITION Mus musculus claudin-18A2.2 mRNA, complete cds, alternatively spliced.

ACCESSION AF349453

VERSION AF349453.1 GI:16224171

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 861)

REFERENCE

AUTHORS Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B.,

Popescu,N.C. and Kimura,S.

claudin-18, a novel downstream target gene for the T/EBP/NKX2.1

homeodomain transcription factor, encodes lung- and

stomach-specific isoforms through alternative splicing

Mol. Cell. Biol. 21 (21), 7380-7390 (2001)

JOURNAL MEDLINE 21470315

PUBMED 11585919

REFERENCE 2 (bases 1 to 861)

AUTHORS Niimi,T. and Kimura,S.

Direct Submission

TITLE

Submitted (15-FEB-2001) Laboratory of Metabolism, National Cancer

Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA

FEATURES

source

1..861

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/note="isolated from mixed background of 129/Sv, C57BL/6, and Black Swiss mice"

54..680

/note="alternatively spliced"

/codon_start=1

/product="claudin-18A2.2"

/protein_id="AAL15638.1"

/db_xref="GI:16224172"

/translation="MSVTACQGLGFVVSLIGFAGIIAATCMDQWSTQDLYNNPVTAVF NYQGLWRSCVRESSGFTECRGYFTLLGLPAMLQAVRALMIVGLVIGILVSIPLAK CIRIGSMDDSAKAKMTLTSGILFIISGICAIIGVSVFANMLVTNFWMSTANMYSGMGG MGGMVQTVQTRYTFGAALFVGWVAGGLTLIGVMMCIACRGLTPDDSK"

ORIGIN

Alignment Scores:

Pred. No.: 2.54e-105 Length: 861
Score: 1147.50 Matches: 217
Percent Similarity: 91.04% Conservative: 27
Best Local Similarity: 80.97% Mismatches: 17
Query Match: 84.56% Indels: 7
DB: 10 Gaps: 2

US-10-063-731-118 (1-261) x AF349453 (1-861)

QY 1 MetSerThrThrCysGlnValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 54 ATGTGGTGACCGCCTGCCAGGGCTTGGGGTTGTGTCACTGATCGGGTTGCGGCG 113
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 114 ATCATTCAGCCACTTGTATGGACCAGGTACCCAGGATTTATACAACACCGCGT 173

| | | | |
|----|-----|---|-----|
| QY | 41 | ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe | 60 |
| DB | 174 | ACCGCTGTATTCAACTACCAAGGGCTATGGCGTTTCATGCGTCCGAGAGAGCTCTGGCTTC | 233 |
| QY | 61 | ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg | 80 |
| DB | 234 | ACCGAGTGGCGAGGCTACTTTCACCCCTGTTGGGGTTGCCAGCCATGCTGCAAGCTGTACGA | 293 |
| QY | 81 | AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla | 100 |
| DB | 294 | GCCCTGATGATCGTGGGCATTGTTCTGGGGGTTCATCGGTATCTCGTGTCCATCTTCGCC | 353 |
| QY | 101 | LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr | 120 |
| DB | 354 | CTGAAGTGCATTGCGATTGGTAGCATGGATGACTCTGCCAAGGCCAAGATGACTCTGACT | 413 |
| QY | 121 | SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla | 140 |
| DB | 414 | TCTGGGATCTTGTTTCATCATCTCCGCGCATCTGTGCAATCATTTGGTGTGTCTGTGTTTGCC | 473 |
| QY | 141 | AsnMetLeuValThrAsnPheTyrMetSerThrAlaAsnMetTyrThr | 157 |
| DB | 474 | AACATGCTGGTGACCAACTTCTGGATGTCCACAGCTAACATGTACAGCGGCATGGCGGC | 533 |
| QY | 158 | MetGlyGlyMetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheVal | 177 |
| DB | 534 | ATGGGTGGCATGGTGCAGACCGTTTCAGACCAGGTACACCTTCGGTGCAGCTCTGTTCGTG | 593 |
| QY | 178 | GlyTyrValAlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArg | 197 |
| DB | 594 | GGCTGGGTGCTGGAGGCCTCACCCCTGATTGGGGAGTGTGATGTGCATCGCCTGCGCT | 653 |
| QY | 198 | GlyLeuAlaProGluGluThrAsn | 213 |
| DB | 654 | GGCCTGACACCCAGATGACAGCAAGTGAGTCTCCCCCTTCAAGCTGTGCTTACCATGCC | 713 |
| QY | 214 | SerGlyHisSerValAlaTyrLysProGlyGlyPheLysAlaSerThrGlyPheGlySer | 233 |
| DB | 714 | TCTGGCCCAAAATGTTGCCTACAGGCTGGAGGCTTTAAGGCCAGCACTGGCTTTGGGTCC | 773 |
| QY | 234 | AsnThrLysAsnLysLysIleTyrAspGlyGlyAlaArgThrGluAspGluValGlnSer | 253 |
| DB | 774 | AACACCAGAAACAAGAAGATCTACGATGGGGGTGCCCCGCACAGAAGACGATGAACAGTCT | 833 |
| QY | 254 | TyrProSerLysHisAspTyrVal | 261 |
| DB | 834 | CATCCTACCAAGTATGACTATGTG | 857 |

| RESULT 14 | LOCUS | BD237993 | 783 bp | DNA | linear | PAT 17-JUL-2003 |
|------------|--|----------|--------|-----|--------|-----------------|
| DEFINITION | Gastric polypeptide ZSIG28. | | | | | |
| ACCESSION | BD237993 | | | | | |
| VERSION | BD237993.1 GI:33047763 | | | | | |
| KEYWORDS | JP 2002524103-A/2. | | | | | |
| SOURCE | synthetic construct | | | | | |
| ORGANISM | artificial sequences. | | | | | |
| REFERENCE | 1 (bases 1 to 783) | | | | | |
| AUTHORS | Sheppard, P.O. and Foley, K.P. | | | | | |
| TITLE | Gastric polypeptide ZSIG28 | | | | | |
| JOURNAL | Patent: JP 2002524103-A 2 06-AUG-2002; | | | | | |
| | ZYMOGENETICS INC | | | | | |
| COMMENT | OS Artificial Sequence | | | | | |
| | PN JP 2002524103-A/2 | | | | | |
| | PD 06-AUG-2002 | | | | | |
| | PF 14-SEP-1999 JP 2000570197 | | | | | |
| | PR 16-SEP-1998 US 09/154444 | | | | | |
| | PI PAUL, O SHEPPARD, KEVIN P FOLEY | | | | | |
| | PC C12N15/09, A61K38/00, A61K39/395, A61K39/395, A61P1/04, A61P1/14, | | | | | |
| | PC A61P3/08, | | | | | |
| | PC A61P5/50, A61P31/04, A61P31/10, A61P35/00, C07K14/47, C07K16/18, PC | | | | | |

```

C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12P21/08,C12Q1/02 PC
,C12Q1/68,G01N33/15,
PC G01N33/50,G01N33/53,G01N33/577,C12N15/00,C12N5/00,A61K37/02 CC
A,Degenerate nucleotide sequence of zsig28 polypeptide CC n =
A,T,C or G
FH Key Location/Qualifiers
FT misc feature (1)..(783).
FEATURES
source
1..783
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Alignment Scores:
Pred. No.: 1.26e-102 Length: 783
Score: 1120.00 Matches: 211
Percent Similarity: 80.84% Conservative: 0
Best Local Similarity: 80.84% Mismatches: 50
Query Match: 82.54% Indels: 0
DB: 6 Gaps: 0

```

QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 721 TAYGAYGGNGGNCNMGACNGARGAYGARGTNCARWSNTAYCCNWSNAARCAYGATAY 780

QY 261 Val 261
Db 781 GTN 783

RESULT 15
BC063845
LOCUS BC063845 2040 bp mRNA linear PRI 12-DEC-2003
DEFINITION Homo sapiens cDNA clone IMAGE:6177081, partial cds.
ACCESSION BC063845
VERSION BC063845.1 GI:39795429
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2040)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettaman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2040)
Strausberg, R.
Direct Submission
Submitted (08-DEC-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 141 Row: a Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7705960.
Location/Qualifiers

JOURNAL PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

FEATURES

source 1. .2040
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6177081"
/tissue_type="Peripheral Nervous System, dorsal root
ganglion"
/clone_lib="Lupski_dorsal_root_ganglion"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:
Pred. No.: 7.64e-101 Length: 2040
Score: 1107.50 Matches: 221
Percent Similarity: 84.67% Conservative: 0
Best Local Similarity: 84.67% Mismatches: 1
Query Match: 81.61% Indels: 40
DB: 9 Gaps: 1

US-10-063-731-118 (1-261) x BC063845 (1-2040)

QY 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 11 ATGTCCACCACCATGCCAAGTGGTGGCTCTCTCTCCATCTCTGGGCTGGCCGGC 70

QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 71 TGCATCGCGCCACCGGATGGACATGTGGAGCACCAGGACCTGTACGACCAACCCCGTC 130

QY 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
Db 131 ACCTCCGTGTCTCCAGTACGAAGGGCTCTGGAGGAGTGGTGGTGGTGGTGGTGGTGGT 190

QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 191 ACCGAATGCAGGCCCTATTTCACCATCTCTGGACTTCCAGCCATGTGCAGGCAGTGCGA 250

QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 251 GCCCTGATGATCGTAGGCATCGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC 310

QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 311 CTGAAATGCATCCGATGGCAGCATGGAGACTCTGCCAAAGCCAACATGACACTGACC 370

QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 371 TCCGGGATCATGTTTCATTGTCTC----- 393

QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 393 ----- 393

QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTyrVal 180
Db 394 -----AGGTACACATTGGTGGCTCTGTTCGTGGGTGGGTC 432

QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 433 GCTGGAGGCCTCACACTAATTGGGGTGTGATGATGTGCATCGCCTCGCGGGGCTGGCA 492

QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 493 CCAGAAAGAAACCACTACAAAGCCGTTTCTTATCATGCCTCGGGGCCACAGTGTTCCTAC 552

QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
Db 553 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTGGTCCACACCAACCAAGAGATA 612

QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 613 TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAAATCTTATCTCTTCCAAAGCACGACTAT 672

Qy 261 Val 261
Db 673 GTG 675

Search completed: September 2, 2004, 21:14:00
Job time : 2837 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 2, 2004, 20:04:09 ; Search time 327 Seconds
(without alignments)
3390.764 Million cell updates/sec

Title: US-10-063-731-118
Perfect score: 1357
Sequence: 1 MSTTTCQVAVFLLSILGLAG.....DGGARTEDEVQSPSKHDYV 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool/US10063731/runat_01092004_155041_18736/app query.fasta_1.455
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10063731 @CGN 1 1 352 @runat_01092004_155041_18736 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002s.*

7: geneseqn2003as.*

8: geneseqn2003bs.*

9: geneseqn2003cs.*

10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1357 | 100.0 | 982 | 3 | AAZ52249 Human sto |
| 2 | 1357 | 100.0 | 1505 | 3 | AAC74775 Human ORF |
| C 3 | 1357 | 100.0 | 1530 | 3 | AAA09120 Clone 322 |
| 4 | 1357 | 100.0 | 1530 | 3 | AAA09116 Clone 322 |
| 5 | 1357 | 100.0 | 1869 | 6 | ABK81817 DNA repre |
| 6 | 1357 | 100.0 | 2108 | 4 | AAF54432 DNA encod |
| 7 | 1357 | 100.0 | 2121 | 3 | AAA37114 Human PRO |
| 8 | 1357 | 100.0 | 2121 | 4 | AAS46102 Human DNA |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| | | | | | | |
|----|------|-------|------|---|----------|--------------------|
| 9 | 1357 | 100.0 | 2121 | 4 | AAF92116 | Aaf92116 Human PRO |
| 10 | 1357 | 100.0 | 2121 | 6 | ABS74436 | AbS74436 Human CDN |
| 11 | 1357 | 100.0 | 2121 | 7 | ABX78705 | Abx78705 Human PRO |
| 12 | 1357 | 100.0 | 2121 | 7 | ACA75677 | Aca75677 Novel hum |
| 13 | 1357 | 100.0 | 2121 | 7 | ACA71157 | Aca71157 Human sec |
| 14 | 1357 | 100.0 | 2121 | 7 | ACC87685 | Acc87685 Human sec |
| 15 | 1357 | 100.0 | 2121 | 7 | ACC87071 | Acc87071 Human sec |
| 16 | 1357 | 100.0 | 2121 | 7 | ACD04244 | ACd04244 Human sec |
| 17 | 1357 | 100.0 | 2121 | 7 | ACA69575 | Aca69575 CDNA enco |
| 18 | 1357 | 100.0 | 2121 | 7 | ACA90420 | Aca90420 Novel hum |
| 19 | 1357 | 100.0 | 2121 | 7 | ACC89527 | Acc89527 Human sec |
| 20 | 1357 | 100.0 | 2121 | 7 | ACA98318 | Aca98318 Novel hum |
| 21 | 1357 | 100.0 | 2121 | 7 | ACA93960 | Aca93960 Human sec |
| 22 | 1357 | 100.0 | 2121 | 7 | ACD15353 | Acd15353 Human sec |
| 23 | 1357 | 100.0 | 2121 | 7 | ACD08940 | Acd08940 Human sec |
| 24 | 1357 | 100.0 | 2121 | 7 | ACC96860 | Acc96860 Human sec |
| 25 | 1357 | 100.0 | 2121 | 7 | ACF15581 | Acf15581 Human sec |
| 26 | 1357 | 100.0 | 2121 | 7 | ACA72948 | Aca72948 Human PRO |
| 27 | 1357 | 100.0 | 2121 | 7 | ACD03120 | Acd03120 Novel hum |
| 28 | 1357 | 100.0 | 2121 | 7 | ACD01935 | Acd01935 Novel hum |
| 29 | 1357 | 100.0 | 2121 | 7 | ACA92127 | Aca92127 Novel hum |
| 30 | 1357 | 100.0 | 2121 | 7 | ACA89552 | Aca89552 CDNA enco |
| 31 | 1357 | 100.0 | 2121 | 7 | ACA73562 | Aca73562 Human sec |
| 32 | 1357 | 100.0 | 2121 | 7 | ACA05877 | Aca05877 Human sec |
| 33 | 1357 | 100.0 | 2121 | 7 | ACA66711 | Aca66711 CDNA enco |
| 34 | 1357 | 100.0 | 2121 | 7 | ACA91222 | Aca91222 Novel hum |
| 35 | 1357 | 100.0 | 2121 | 7 | ACD81599 | Acd81599 Human CDN |
| 36 | 1357 | 100.0 | 2121 | 7 | ACF20286 | Acf20286 Human sec |
| 37 | 1357 | 100.0 | 2121 | 7 | ACF19672 | Acf19672 Human sec |
| 38 | 1357 | 100.0 | 2121 | 7 | ACD21960 | Acd21960 Human sec |
| 39 | 1357 | 100.0 | 2121 | 7 | ACF13125 | Acf13125 Human sec |
| 40 | 1357 | 100.0 | 2121 | 7 | ACD25228 | Acd25228 Human sec |
| 41 | 1357 | 100.0 | 2121 | 7 | ACF00277 | Acf00277 Human sec |
| 42 | 1357 | 100.0 | 2121 | 7 | ACA60421 | Aca60421 Novel hum |
| 43 | 1357 | 100.0 | 2121 | 7 | ACA72334 | Aca72334 Novel hum |
| 44 | 1357 | 100.0 | 2121 | 7 | ACD04858 | Acd04858 Novel hum |
| 45 | 1357 | 100.0 | 2121 | 7 | ACD18319 | Acd18319 Human sec |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| AAZ52249 | |
| ID | AAZ52249 standard; DNA; 982 BP. |
| XX | |
| AC | AAZ52249; |
| XX | |
| DT | 18-JUL-2000 (first entry) |
| XX | |
| DE | Human stomach protein zsig28 DNA. |
| XX | |
| KW | Human; stomach; zsig28 protein; chromosome 3q22.1-3q22.2; gene therapy; |
| KW | claudin; oligodendrocyte-specific protein; OSP; apoptosis; RVP.1; |
| KW | rat androgen-withdrawal apoptosis protein; growth factor receptor; |
| KW | cell-cell signalling molecule; cytostatic; antibacterial; food poisoning; |
| KW | Botulism; diarrhoea; inflammation; cramping; cancer; gastric ulcer; |
| KW | diagnosis; prevention; treatment; ds. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FH | Key |
| FT | CDS |
| FT | Location/Qualifiers |
| FT | 70..855 |
| FT | /*tag= a |
| FT | /product= "zsig28 protein" |
| FT | 70..138 |
| FT | /*tag= b |
| FT | 139..852 |
| FT | /*tag= c |
| FT | /product= "Mature zsig28" |
| XX | |
| PN | WO200015659-A2. |
| XX | |
| PD | 23-MAR-2000. |

14-SEP-1999; 99WO-US021023.
 16-SEP-1998; 98US-00154444.
 (ZYMO) ZYMOGENETICS INC.
 Sheppard PO, Foley KP;
 WPI; 2000-271379/23.
 P-PSDB; AAY70675.

New isolated polynucleotide encoding a stomach zsig28 polypeptide used for diagnosis, prevention and treatment of stomach disorders caused by bacteria, gastric ulcers or cancer.

Claim 2; Page 111-113; 121pp; English.

The present sequence is a stomach protein zsig28 encoding DNA located at 3q22.1-3q22.2 region of human chromosome 3 and isolated from human lung library. The zsig28 protein shows homology to a diverse family of receptor proteins containing e.g. human Claudin 1 and 2, human and murine oligodendrocyte-specific protein (OSP) and rat androgen-withdrawal apoptosis protein RVP.1. It is thought to be a cell-cell signalling molecule, a growth factor receptor or extracellular matrix associated protein with growth factor hormone activity and may be involved in an apoptotic cellular pathway. The protein may act as an anti-microbial agent and may bind toxins produced by bacteria which cause food poisoning, Botulism, severe diarrhoea, inflammation and cramping. zsig28 agonists are useful for promoting apoptosis in cells over-expressing zsig28 e.g. in cancer cells. They are also useful for stimulating cell growth or differentiation. Altered levels of zsig28 protein in a test sample such as saliva, serum, sweat or biopsy can be monitored as an indication of digestive function, gastric ulcer or cancer. zsig28 expression can be used as a differentiation marker to determine the stage of tumour or cell maturity, particularly in epithelial cells. Polynucleotides encoding zsig28 can be used in gene therapy applications to increase or inhibit zsig28 activity

Sequence 982 BP; 218 A; 275 C; 271 G; 218 T; 0 U; 0 Other;

| Alignment Scores: | | | | |
|------------------------|----------|---------------|-----|--|
| Pred. No.: | 3.4e-149 | Length: | 982 | |
| Score: | 1357.00 | Matches: | 261 | |
| Percent Similarity: | 100.00% | Conservative: | 0 | |
| Best Local Similarity: | 100.00% | Mismatches: | 0 | |
| Query Match: | 100.00% | Indels: | 0 | |
| DB: | 3 | Gaps: | 0 | |

US-10-063-731-118 (1-261) x AAZ52249 (1-982)

| | | | |
|----|-----|--|-----|
| QY | 1 | MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly | 20 |
| Db | 70 | ATGTCACACACACATGCCAAGTGTGGCGTTCTCTCTGTCCATCTGGGGTGGCGGC | 129 |
| QY | 21 | CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal | 40 |
| Db | 130 | TGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACACCCCGTC | 189 |
| QY | 41 | ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe | 60 |
| Db | 190 | ACCTCCGTGTCCAGTACGAAGGGCTCTGGAGGAGCTGCTGAGGCAGAGTTCAGGCTTC | 249 |
| QY | 61 | ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg | 80 |
| Db | 250 | ACCGAATGCAGGCCCTATTTCCACCATCTGGGACTTCCAGCCATGCTGCAGGACGTGCGA | 309 |
| QY | 81 | AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla | 100 |
| Db | 310 | GCCCTGATGATCGTAGGCATCGTCTCGGGTGCCATTGGCCCTCTGGTATCCATCTTTGCC | 369 |
| QY | 101 | LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr | 120 |

XX Shimkets RA, Leach M;
PI
XX
DR WPI; 2000-602362/57.
DR P-PSDB; AAB40566.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
PS Claim 5; Page 771; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 1505 BP; 367 A; 398 C; 337 G; 403 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.34e-149 Length: 1505
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-063-731-118 (1-261) x AAC74775 (1-1505)

Qy 1 MetSerThrThrCysGlnValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db
Qy 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 70 TGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAAACCCCGTC 129
Qy 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
Db 130 ACCTCGGTGTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTC 189
Qy 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 190 ACCGAATGCAGGCCCTATTTCACCATCTCGGACTTCAGCCATGTGCGAGGCAGTGCGA 249
Qy 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 250 GGCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCCTGGTATCCATCTTTGCC 309
Qy 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 310 CTGAAATGCATCCGATGGCAGCATGGAGGACTCTGCCAAGCCACATGACACATGACC 369
Qy 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 370 TCCGGGATCATGTTTCATGTCTCAGGCTTTTGCAATTGCTGGAGTGTCTGTGTTTGGC 429

Qy 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 430 AACATGCTGGTGAATAACTTCTGGATGTCCACAGCTAACATGTACACGGCATGGGTGG 489
Qy 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 490 ATGGTGCAGACTGTTCAGACCAGGTACACATTGGTGGGCTCTGTTCGTGGCTGGGTC 549
Qy 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 550 GCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCTGCCGGGCTGGCA 609
Qy 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 610 CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTGCCTAC 669
Qy 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
Db 670 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTGGGTCCAACACCAAAAAACAAGAAGATA 729
Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 730 TACGATGGAGGTGCCCGCACAGAGGACGAGGTACATCTTATCTCTCCAAGCAGACTAT 789
Qy 261 Val 261
Db 790 GTG 792
RESULT 3
AAA09120/c
ID AAA09120 standard; cDNA; 1530 BP.
XX
AC AAA09120;
XX
DT 10-AUG-2000 (first entry)
DE Clone 3224646 cDNA complement.
XX
KW Clone 3224646; claudin; homologue; cytostatic; anti-HIV;
KW immunosuppressive; antiallergic; antiinfective; antiinflammatory;
KW antiarthritic; antiarteriosclerotic; vasotropic; neuroprotective;
KW nootropic; dermatological; tranquilizer; vulnery; ss.
XX
OS Homo sapiens.
XX
PN WO200020447-A2.
XX
PD 13-APR-2000.
XX
PF 06-OCT-1999; 99WO-US023294.
XX
PR 06-OCT-1998; 98US-0103195P.
PR 05-OCT-1999; 99US-00412231.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA;
XX
DR WPI; 2000-303741/26.
XX
PT Nucleic acids encoding polypeptides with syncline-like, claudin-like or
PT cytokine-like activity, useful for treating diseases including cancer,
PT Alzheimer's and atherosclerosis.
XX
PS Claim 12; Fig 3B; 118pp; English.
XX
CC Clone 3223867 encodes a polypeptide that has homology to claudin-1, which
CC is an integral membrane protein found in tight junctions. The sequences
CC are useful for treatment of diseases such as cancer, immune disorders,
CC autoimmune disease, acquired immune deficiency syndrome (AIDS),
CC transplant rejection, allergy, infection by a pathological agent or
CC organism, inflammatory disorders, arthritis, a haematopoietic disorder, a
CC skin disorder, atherosclerosis, restenosis, a neurological disease,

| | |
|----|--|
| CC | Alzheimer's disease, trauma, spinal cord injury and skeletal disorders |
| XX | |
| SQ | Sequence 1530 BP; 375 A; 381 C; 408 G; 366 T; 0 U; 0 Other; |

| | | |
|------------------------|---------------|------|
| Alignment Scores: | | |
| Pred. No.: | 6.5e-149 | 1530 |
| Score: | 1357.00 | 261 |
| Percent Similarity: | 100.00% | 0 |
| Best Local Similarity: | 100.00% | 0 |
| Query Match: | 100.00% | 0 |
| DB: | 3 | 0 |
| | Length: | |
| | Matches: | |
| | Conservative: | |
| | Mismatches: | |
| | Indels: | |
| | Gaps: | |

US-10-063-731-118 (1-261) x AAA09120 (1-1530)

| | | | |
|----|------|--|------|
| QY | 1 | MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly | 20 |
| DB | 1230 | ATGTCCACCACCATGCCAAGTGGTGGCGTTCCCTCTGTCCATCTCTGGGCTGGCCGGC | 1171 |
| QY | 21 | CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal | 40 |
| DB | 1170 | TGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACACCCCGTC | 1111 |
| QY | 41 | ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe | 60 |
| DB | 1110 | ACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTC | 1051 |
| QY | 61 | ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg | 80 |
| DB | 1050 | ACCGAATGCAGGCCCTATTTCACCATCCTGGGACCTCCAGCCATGCTGCAGGCAGTGGCA | 991 |
| QY | 81 | AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla | 100 |
| DB | 990 | GCCCTGATGATCGTAGGCATCGTCTCTGGGTGCCATTGGCTCTCTGGTATCCATCTTTGCC | 931 |
| QY | 101 | LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAasnMetThrLeuThr | 120 |
| DB | 930 | CTGAAATGCATCCGCATTTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACACTGACC | 871 |
| QY | 121 | SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla | 140 |
| DB | 870 | TCCGGGATCATGTTCAATGTCTCAGGTCTTTGTGCAATGCTGGAGTGTCTGTGTTTGGC | 811 |
| QY | 141 | AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly | 160 |
| DB | 810 | AACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGGG | 751 |
| QY | 161 | MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal | 180 |
| DB | 750 | ATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGCGCTCTGTTCTGGTGGGCTGGGTC | 691 |
| QY | 181 | AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla | 200 |
| DB | 690 | GCTGGAGGCCTCACATAATTGGGGTGTGATGATGTGCATCGCCTGCCGGGCCCTGGCA | 631 |
| QY | 201 | ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr | 220 |
| DB | 630 | CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTGGCTTAC | 571 |
| QY | 221 | LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle | 240 |
| DB | 570 | AAGCCTGGAGGCTCAAGGCCAGCACTGGCTTTGGTCCACACCCAAACCAAGAGATA | 511 |
| QY | 241 | TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr | 260 |
| DB | 510 | TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAACTTATCTTATCTTCCAAAGCACGACTAT | 451 |
| QY | 261 | Val | 261 |
| DB | 450 | GTG | 448 |

RESULT 4
AAA09116
ID AAA09116 standard; cDNA; 1530 BP.

| | | | |
|----|--------------------|--|--|
| XX | AAA09116; | | |
| AC | | | |
| XX | | | |
| DT | 10-AUG-2000 | (first entry) | |
| XX | | | |
| DE | Clone 3224646 | cdNA, encodes claudin homologue. | |
| XX | | | |
| KW | Clone 3224646; | claudin; homologue; cytostatic; anti-HIV; | |
| KW | immunosuppressive; | antiallergic; antiinfective; antiinflammatory; | |
| KW | antiarthritic; | antiarteriosclerotic; vasotropic; neuroprotective; | |
| KW | nootropic; | dermatological; tranquilizer; vulnerary; ss. | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| FH | Key | Location/Qualifiers | |
| FT | CDS | 301..1086 | |
| FT | | /*tag= a | |
| FT | sig_peptide | 301..369 | |
| FT | | /*tag= b | |
| FT | mat_peptide | 370..1083 | |
| FT | | /*tag= c | |

| | | |
|------------------------|----------|-----------------|
| Alignment Scores: | | |
| Pred. No.: | 6.5e-149 | Length: 1530 |
| Score: | 1357.00 | Matches: 261 |
| Percent Similarity: | 100.00% | Conservative: 0 |
| Best Local Similarity: | 100.00% | Mismatches: 0 |
| Query Match: | 100.00% | Indels: 0 |
| DB: | 3 | Gaps: 0 |

US-10-063-731-118 (1-261) x AAA09116 (1-1530)

[illegible]

QY 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
Db 421 ACCTCCGTGTTCCAGTACGAAGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTTCAGGCTTC 480
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 481 ACCGAATGCAGGCCCTATTTCCACCATCCTGGACTTCAGCCATGCTGCAGGCAGTGCGA 540
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 541 GCCCTGATGATCGTAGGCATCGCTGGTGCCATTGGCTCCTGGTATCCATCTTTGCC 600
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 601 CTGAATGCATCCGCAATTGGCAGCATGGAGGACTTGCCAAAGCCAAACATGACACTGACC 660
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 661 TCCGGGATCATGTTCAATTGCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 720
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 721 AACATGCTGGTGACTAACTTCTGGATGTCACAGCTAACATGTACACCGGCATGGGTGG 780
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 781 ATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGGCTCTGTTTCGTGGGCTGGGTC 840
QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 841 GCTGGAGGCCTCACACTAATTGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCA 900
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 901 CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTTCAGGCCACAGTGTGGCTAC 960
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
Db 961 AAGCCTGGAGGCTTCAAGCCAGCACTGGCTTTGGGTCCAACACCAAAACAAGAGATA 1020
QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 1021 TACGATGGAGGTGCCCCGACAGAGGACGAGGTACAATCTTATCCTTCCAGCAGCACTAT 1080
QY 261 Val 261
Db 1081 GTG 1083
RESULT 5
ABK81817
ID ABK81817 standard; DNA; 1869 BP.
XX
AC ABK81817;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA representing lung specific gene #3.
XX
KW Lung specific gene; gene therapy; vaccine; lung cancer; cancer staging;
KW cancer monitoring; cancer diagnosis; imaging lung cancer; metastases;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200218576-A2.
XX
PD 07-MAR-2002.
XX
XX 27-AUG-2001; 2001WO-US025684.
PF
XX 28-AUG-2000; 2000US-0228378P.
PR
XX (DIAD-) DIADEXUS INC.
PA

XX
PI
XX
DR
XX
PT
PT
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
XX
SQ
Alignment Scores:
Pred. No.: 8.71e-149 Length: 1869
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-063-731-118 (1-261) x ABK81817 (1-1869)
QY 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 83 ATGTCCACCACCACATGCCAAGTGTGGCGTTCTCTCTGTCCATCCTGGGCTGGCCGGC 142
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 143 TGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAAACCCGTC 202
QY 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
Db 203 ACCTCCGTGTTCCAGTACGAAGGCTCTGGAGAGCTCGGTGAGGCAGAGTTTCAGGCTTC 262
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 263 ACCGAATGCAGGCCCTATTTCCACATCCTGGGACTTCAGCCATGCTGCAGGCAGTGCGA 322
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 323 GCCCTGATGATCGTAGGCATCGTCTGGTGGTGCATTTGGCTCCTCGGTATCCATCTTTGCC 382
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 383 CTGAATGCATCCGCAATTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACACTGACC 442
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 443 TCCGGGATCATGTTCAATTGCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 502
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 503 AACATGCTGGTGACTAACTTCTGGATGTCACAGCTAACATGTACACCGGCATGGGTGGG 562
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 563 ATGGTGCAGACTGTTCCAGACCAGGTACACATTTGGTGGGCTCTGTTCTGTTGGGCTGGGTC 622
QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 623 GCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCTGCCGGGGCCTGGCA 682
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220

Chen S, Macina RA, Sun Y, Recipon H;

WPI; 2002-434904/46.

New lung specific genes and their encoded proteins, useful in gene therapy or as a vaccine for treating lung cancer, as well as for measuring metastases of lung cancer, or staging, monitoring, diagnosing or imaging lung cancer.

Claim 1; Page 160-161; 206pp; English.

The invention describes a new lung specific gene and it's variants. The lung specific gene proteins and genes are useful in gene therapy or as a vaccine for treating lung cancer. Lung specific genes are also useful for staging, monitoring, diagnosing or imaging lung cancer, as well as for measuring metastases of lung cancer. This sequence represents a lung specific gene described in the invention

Sequence 1869 BP; 485 A; 457 C; 412 G; 515 T; 0 U; 0 Other;

| | | | |
|-------------------|---|--|------|
| Db | 683 | CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTAC | 742 |
| QY | 221 | LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle | 240 |
| Db | 743 | AAGCCTGGAGGCTTCAAGCCAGCACTGGCTTTGGGTCCAACACCAAAACAAGAAGATA | 802 |
| QY | 241 | TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr | 260 |
| Db | 803 | TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAGCAGACTAT | 862 |
| QY | 261 | Val 261 | |
| Db | 863 | GTG 865 | |
| RESULT 6 | | | |
| ID | AAF54432 | standard; DNA; 2108 BP. | |
| XX | | | |
| AC | AAF54432; | | |
| XX | | | |
| DT | 02-APR-2001 | (first entry) | |
| XX | | | |
| DE | DNA encoding protein of the invention #91. | | |
| XX | | | |
| KW | Secreted; transmembrane; gene therapy; ss. | | |
| XX | | | |
| OS | Unidentified. | | |
| XX | | | |
| PN | WO200078961-A1. | | |
| XX | | | |
| PD | 28-DEC-2000. | | |
| XX | | | |
| PF | 18-FEB-2000; 2000WO-US004342. | | |
| XX | | | |
| PR | 23-JUN-1999; 99US-0141037P. | | |
| PR | 20-JUL-1999; 99US-0144758P. | | |
| PR | 26-JUL-1999; 99US-0145698P. | | |
| PR | 01-SEP-1999; 99WO-US020111. | | |
| PR | 29-OCT-1999; 99US-0162506P. | | |
| PR | 30-NOV-1999; 99WO-US028313. | | |
| PR | 02-DEC-1999; 99WO-US028551. | | |
| PR | 16-DEC-1999; 99WO-US030095. | | |
| PR | 05-JAN-2000; 2000WO-US000219. | | |
| PR | 06-JAN-2000; 2000WO-US000376. | | |
| XX | | | |
| PA | (GETH) GENENTECH INC. | | |
| XX | | | |
| PI | Baker KP, Botatein D, Desnoyers L, Eaton DL, Ferrara N, Fong S; | | |
| PI | Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ; | | |
| PI | Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK; | | |
| PI | Williams PM, Wood WI; | | |
| XX | | | |
| DR | WPI; 2001-071395/08. | | |
| XX | | | |
| PT | Secreted and transmembrane proteins and nucleic acids designated PRO, | | |
| PT | useful as hybridization probes, in chromosome and gene mapping and gene | | |
| PT | therapy. | | |
| XX | | | |
| PS | Claim 2; Fig 181; 787pp; English. | | |
| XX | | | |
| CC | The present invention relates to secreted and transmembrane proteins. | | |
| CC | These proteins and the DNA encoding them may be used as hybridization | | |
| CC | probes, in chromosome and gene mapping and in the generation of anti- | | |
| CC | sense RNA and DNA. They may also be used to generate either | | |
| CC | transgenic animals or knockout animals which are in turn useful for | | |
| CC | development and screening of therapeutically useful reagents. The nucleic | | |
| CC | acids may also be used in gene therapy | | |
| XX | | | |
| SQ | Sequence 2108 BP; 564 A; 514 C; 487 G; 543 T; 0 U; 0 Other; | | |
| Alignment Scores: | | | |
| Pred. No.: | 1.04e-148 | Length: | 2108 |
| Score: | 1357.00 | Matches: | 261 |

| | | | |
|---|------------------------|--|-----|
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |
| US-10-063-731-118 (1-261) x AAF54432 (1-2108) | | | |
| QY | 1 | MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly | 20 |
| Db | 90 | ATGTCCACCACCACCATGCCAAGTGGTGGCGTCTCTGTCCATCCTGGGGTGGCCGGC | 149 |
| QY | 21 | CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal | 40 |
| Db | 150 | TGCATCGGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACACCCCGTC | 209 |
| QY | 41 | ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe | 60 |
| Db | 210 | ACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGAGCTCGTGAGGCAGAGTTCAGGCTTC | 269 |
| QY | 61 | ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg | 80 |
| Db | 270 | ACCGAATGCAGGCCCTATTTCACCATCTCTGGACTTCCAGCCATGCTGAGGCAGTGC | 329 |
| QY | 81 | AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla | 100 |
| Db | 330 | GCCTGATGATCGTAGGCATCGTCTGGTGCCATTTGGCTCTCTGTATCATCTTTGCC | 389 |
| QY | 101 | LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr | 120 |
| Db | 390 | CTGAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAGCCAACATGACATGACC | 449 |
| QY | 121 | SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla | 140 |
| Db | 450 | TCCGGATCATGTTCTCATTTCTCAGGTCTTTTGCAATTGCTGGAGTGTCTGTGTTGCC | 509 |
| QY | 141 | AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly | 160 |
| Db | 510 | AACATGCTGGTGACTAACTTCTGGATGTCCACAGTAACATGTACCCGGCATGGGTGG | 569 |
| QY | 161 | MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal | 180 |
| Db | 570 | ATGGTGCACTGTTTCAGACCAAGGTACACATTTGGTGGGCTCTGTTCGTGGGCTGGGTC | 629 |
| QY | 181 | AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla | 200 |
| Db | 630 | GCTGGAGGCCTCACACTAATTTGGGGTGTGATGATGTGCATCGCTGCCGGGCTGGCA | 689 |
| QY | 201 | ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr | 220 |
| Db | 690 | CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTGCCTAC | 749 |
| QY | 221 | LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle | 240 |
| Db | 750 | AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAACAAGAAGATA | 809 |
| QY | 241 | TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr | 260 |
| Db | 810 | TACGATGGAGGTGCCCCGACAGAGGACGAGGTACAATCTTATCCTTCCAAGCAGACTAT | 869 |
| QY | 261 | Val 261 | |
| Db | 870 | GTG 872 | |
| RESULT 7 | | | |
| AAA37114 | | | |
| ID | AAA37114 | standard; cDNA; 2121 BP. | |
| XX | | | |
| AC | AAA37114; | | |
| XX | | | |
| DT | 08-AUG-2000 | (first entry) | |
| XX | | | |
| DE | Human PRO1572 (UNQ778) | cDNA sequence SEQ ID NO:325. | |
| XX | | | |

KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening; ss.

XX Homo sapiens.

XX WO200012708-A2.

XX 09-MAR-2000.

XX 01-SEP-1999; 99WO-US020111.

PR 01-SEP-1998; 98US-0098716P.

PR 01-SEP-1998; 98US-0098749P.

PR 01-SEP-1998; 98US-0098750P.

PR 02-SEP-1998; 98US-0098803P.

PR 02-SEP-1998; 98US-0098821P.

PR 02-SEP-1998; 98US-0098843P.

PR 09-SEP-1998; 98US-0099536P.

PR 09-SEP-1998; 98US-0099596P.

PR 09-SEP-1998; 98US-0099598P.

PR 09-SEP-1998; 98US-0099602P.

PR 09-SEP-1998; 98US-0099642P.

PR 10-SEP-1998; 98US-0099741P.

PR 10-SEP-1998; 98US-0099754P.

PR 10-SEP-1998; 98US-0099763P.

PR 10-SEP-1998; 98US-0099792P.

PR 10-SEP-1998; 98US-0099808P.

PR 10-SEP-1998; 98US-0099812P.

PR 10-SEP-1998; 98US-0099815P.

PR 10-SEP-1998; 98US-0099816P.

PR 15-SEP-1998; 98US-0100385P.

PR 15-SEP-1998; 98US-0100388P.

PR 15-SEP-1998; 98US-0100390P.

PR 16-SEP-1998; 98US-0100584P.

PR 16-SEP-1998; 98US-0100627P.

PR 16-SEP-1998; 98US-0100661P.

PR 16-SEP-1998; 98US-0100662P.

PR 16-SEP-1998; 98US-0100664P.

PR 17-SEP-1998; 98US-0100683P.

PR 17-SEP-1998; 98US-0100684P.

PR 17-SEP-1998; 98US-0100710P.

PR 17-SEP-1998; 98US-0100711P.

PR 17-SEP-1998; 98US-0100919P.

PR 17-SEP-1998; 98US-0100930P.

PR 18-SEP-1998; 98US-0100848P.

PR 18-SEP-1998; 98US-0100849P.

PR 18-SEP-1998; 98US-0101014P.

PR 18-SEP-1998; 98US-0101068P.

PR 18-SEP-1998; 98US-0101071P.

PR 22-SEP-1998; 98US-0101279P.

PR 23-SEP-1998; 98US-0101471P.

PR 23-SEP-1998; 98US-0101472P.

PR 23-SEP-1998; 98US-0101474P.

PR 23-SEP-1998; 98US-0101475P.

PR 23-SEP-1998; 98US-0101476P.

PR 23-SEP-1998; 98US-0101477P.

PR 23-SEP-1998; 98US-0101479P.

PR 24-SEP-1998; 98US-0101738P.

PR 24-SEP-1998; 98US-0101741P.

PR 24-SEP-1998; 98US-0101743P.

PR 24-SEP-1998; 98US-0101915P.

PR 24-SEP-1998; 98US-0101916P.

PR 29-SEP-1998; 98US-0102207P.

PR 29-SEP-1998; 98US-0102240P.

PR 29-SEP-1998; 98US-0102307P.

PR 29-SEP-1998; 98US-0102330P.

PR 29-SEP-1998; 98US-0102331P.

PR 30-SEP-1998; 98US-0102484P.

PR 30-SEP-1998; 98US-0102487P.

PR 30-SEP-1998; 98US-0102570P.

PR 30-SEP-1998; 98US-0102571P.

PR 01-OCT-1998; 98US-0102684P.

PR 01-OCT-1998; 98US-0102687P.

PR 02-OCT-1998; 98US-0102965P.

PR 06-OCT-1998; 98US-0103258P.

PR 06-OCT-1998; 98US-0103449P.

PR 07-OCT-1998; 98US-0103314P.

PR 07-OCT-1998; 98US-0103315P.

PR 07-OCT-1998; 98US-0103328P.

PR 07-OCT-1998; 98US-0103395P.

PR 07-OCT-1998; 98US-0103396P.

PR 07-OCT-1998; 98US-0103401P.

PR 08-OCT-1998; 98US-0103633P.

PR 08-OCT-1998; 98US-0103678P.

PR 08-OCT-1998; 98US-0103679P.

PR 14-OCT-1998; 98US-0103711P.

PR 14-OCT-1998; 98US-0104257P.

PR 20-OCT-1998; 98US-0104987P.

PR 20-OCT-1998; 98US-0105000P.

PR 20-OCT-1998; 98US-0105002P.

PR 21-OCT-1998; 98US-0105104P.

PR 22-OCT-1998; 98US-0105169P.

PR 22-OCT-1998; 98US-0105266P.

PR 26-OCT-1998; 98US-0105693P.

PR 26-OCT-1998; 98US-0105694P.

PR 27-OCT-1998; 98US-0105807P.

PR 27-OCT-1998; 98US-0105881P.

PR 27-OCT-1998; 98US-0105882P.

PR 27-OCT-1998; 98US-0106062P.

PR 28-OCT-1998; 98US-0106023P.

PR 28-OCT-1998; 98US-0106029P.

PR 28-OCT-1998; 98US-0106030P.

PR 28-OCT-1998; 98US-0106032P.

PR 28-OCT-1998; 98US-0106033P.

PR 28-OCT-1998; 98US-0106178P.

PR 29-OCT-1998; 98US-0106248P.

PR 29-OCT-1998; 98US-0106384P.

PR 29-OCT-1998; 98US-0108500P.

PR 30-OCT-1998; 98US-0106464P.

PR 03-NOV-1998; 98US-0106856P.

PR 03-NOV-1998; 98US-0106902P.

PR 03-NOV-1998; 98US-0106905P.

PR 03-NOV-1998; 98US-0106919P.

PR 03-NOV-1998; 98US-0106932P.

PR 03-NOV-1998; 98US-0106934P.

PR 10-NOV-1998; 98US-0107783P.

PR 17-NOV-1998; 98US-0108775P.

PR 17-NOV-1998; 98US-0108779P.

PR 17-NOV-1998; 98US-0108787P.

PR 17-NOV-1998; 98US-0108788P.

PR 17-NOV-1998; 98US-0108801P.

PR 17-NOV-1998; 98US-0108802P.

PR 17-NOV-1998; 98US-0108806P.

PR 17-NOV-1998; 98US-0108807P.

PR 17-NOV-1998; 98US-0108867P.

PR 17-NOV-1998; 98US-0108925P.

PR 18-NOV-1998; 98US-0108848P.

PR 18-NOV-1998; 98US-0108849P.

PR 18-NOV-1998; 98US-0108850P.

PR 18-NOV-1998; 98US-0108851P.

PR 18-NOV-1998; 98US-0108852P.

PR 18-NOV-1998; 98US-0108858P.

PR 18-NOV-1998; 98US-0108904P.

XX (GETH) GENENTECH INC.

PA Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;

XX WPI; 2000-237871/20.

XX P-PSDB; AAY99432.

DR New mammalian DNA sequences encoding transmembrane, receptor or secreted

XX PRO polypeptides, useful for screening of potential peptide or small

PT molecule inhibitors of the relevant receptor/ligand interactions.

XX Claim 2; Fig 185; 773pp; English.

PS

XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
CC receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
CC transmembrane and receptor PRO proteins can be used for screening of
CC potential peptide or small molecule inhibitors of the relevant
CC receptor/ligand interactions. The polypeptides and nucleotide sequences
CC encoding then have various industrial applications, including uses as
CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR
CC primers and hybridisation probes used in the isolation of the PRO
CC polypeptides from the present invention
XX
SQ Sequence 2121 BP; 569 A; 516 C; 487 G; 549 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.05e-148 Length: 2121
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-063-731-118 (1-261) x AAA37114 (1-2121)

QY 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 90 ATGTCCACCACCATGCCAAGTGGTGGCGTTCCTCTGTCCATCTCTGGGCTGGCCGGC 149
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 150 TGCATCGCGGCCACCGGATGGACATGTGGACACCCAGGACCTGTACGACAAACCCCGTC 209
QY 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
Db 210 ACCTCCGTGTTCAGTACGAAGGGCTCTGGAGAGTGGTGGAGGACAGTTCAGGCTTC 269
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 270 ACCGAATGCAGGCCCTATTTCACCATCTCGGACTTCCAGCCATGTGTCAGGCAGTGGCA 329
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 330 GCCCTGATGATCGTAGGCATCGTCTGGTGCCATTTGGCTCTCTGATCCATCTTTGCC 389
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 390 CTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAGCCACATGACACTGACC 449
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 450 TCCGGGATCATGTTTCATGTCTCAGGTCITTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 509
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 510 AACATGCTGGTGACTTAACCTCTGGATGTCACAGCTAACATGTACACCGGCATGGTGGG 569
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTyrVal 180
Db 570 ATGGTGCAGACTGTTACAGACCAGGTACACATTTGGTGGCTCTGTTCGTGGCTGGGTC 629
QY 181 AlaGlyGlyLeuThrLeuIleGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 630 GCTGGAGGCCTCACACTAATTGGGGTGTGATGATGTGATCGCTCGCGGCGCTGGCA 689
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 690 CCAGAAGAAACCACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTTCCTAC 749
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
Db 750 AAGCCTGGAGGCTTCAAGCCAGCACTGGCTTTGGTCCACACCAAAACAAGAGATA 809
QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260

Db 810 TAGATGGAGGTGCCCCGACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGACTAT 869
QY 261 Val 261
|||
Db 870 GTG 872
RESULT 8
AAS46102
ID AAS46102 standard; cDNA; 2121 BP.
XX
AC AAS46102;
XX 18-DEC-2001 (first entry)
XX Human DNA encoding PRO polypeptide sequence #178.
DE PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
KW PCR primer.
XX
OS Homo sapiens.
XX WO200168848-A2.
XX 20-SEP-2001.
PD
XX 28-FEB-2001; 2001WO-US006520.
PF
XX 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 03-MAR-2000; 2000US-0187202P.
PR 06-MAR-2000; 2000US-0186968P.
PR 14-MAR-2000; 2000US-0189320P.
PR 14-MAR-2000; 2000US-0189328P.
PR 15-MAR-2000; 2000WO-US006884.
PR 21-MAR-2000; 2000US-0190828P.
PR 21-MAR-2000; 2000US-0191007P.
PR 21-MAR-2000; 2000US-0191048P.
PR 21-MAR-2000; 2000US-0191314P.
PR 28-MAR-2000; 2000US-0192655P.
PR 29-MAR-2000; 2000US-0193032P.
PR 29-MAR-2000; 2000US-0193053P.
PR 30-MAR-2000; 2000WO-US008439.
PR 04-APR-2000; 2000US-0194449P.
PR 04-APR-2000; 2000US-0194647P.
PR 11-APR-2000; 2000US-0195975P.
PR 11-APR-2000; 2000US-0196000P.
PR 11-APR-2000; 2000US-0196187P.
PR 11-APR-2000; 2000US-0196690P.
PR 11-APR-2000; 2000US-0196820P.
PR 18-APR-2000; 2000US-0198121P.
PR 18-APR-2000; 2000US-0198585P.
PR 25-APR-2000; 2000US-0199397P.
PR 25-APR-2000; 2000US-0199550P.
PR 25-APR-2000; 2000US-0199654P.
PR 03-MAY-2000; 2000US-0201516P.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 05-JUN-2000; 2000US-0209832P.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.
XX
PA (GETH) GENENTECH INC.
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;

PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2001-602746/68.
DR P-PSDB; AAU29201.
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumors, such as prostate and breast tumors, in mammals and to
PT screen for modulators of the compounds.
XX Claim 2; Fig 355; 774pp; English.
PS Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR
XX primers for PRO polypeptides of the invention. The sequences of the
CC invention can be used to detect the presence of a tumour in a mammal by
CC comparing the level of expression of a PRO polypeptide in a test sample
CC of cells from the animal and a control sample of normal cells, whereby a
CC higher level of expression in the test sample indicates the presence of a
CC tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,
CC pigs, goats and rabbits but are preferably human. The polypeptides can be
CC used to stimulate tumour necrosis factor (TNF) alpha release from human
CC blood, when contacted with it. A specific polypeptide can be used to
CC stimulate the proliferation or differentiation of chondrocyte cells. The
CC PRO proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders
XX
SQ Sequence 2121 BP; 569 A; 516 C; 487 G; 549 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,05e-148 Length: 2121
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-063-731-118 (1-261) x AAS46102 (1-2121)

QY 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
DB 90 ATGTCACCCACCATGCCAAGTGGTGGCGTTCCTCTGTCCTCCTGCGGCTGGCCGGC 149

QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
DB 150 TGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACAAACCCGTC 209

QY 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
DB 210 ACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGTCCGTGAGGCAGAGTTCAGGCTTC 269

QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
DB 270 ACCGAATGCAGGCCCTATTTCACCATCCTGGACTTCCAGCCATGCTGCAGGCAGTCCGA 329

QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
DB 330 GCCCTGATGATCGTAGGCATCGTCTGGTGCCATTGGCCTCCTGGTATCCATCTTTGCC 389

QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
DB 390 CTGAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCACACATGACACTGACC 449

QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
DB 450 TCCGGGATCATGTTCAATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 509

QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
DB 510 AACATGCTGGTGACTAACTTCTGGATGTCCAGACTAACATGTACACCGGCATGGGTGGG 569

QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180

DB 570 ATGGTCAGACTGTTTCAGACCAGGTACACATTGGTGGGCTCTGTTCTGGGGCTGGGTC 629
QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
DB 630 GCTGGAGGCCTCACACTAATTGGGGTGTGATGATGTCATCGCCTGCCGGGCTGGCA 689
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
DB 690 CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTAC 749
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
DB 750 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAACAAGAAGATA 809
QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
DB 810 TAGCATGGAGGTGCCCGCACAGAGGACGAGGTACATCTTATCCTTCCAGCAGCACTAT 869
QY 261 Val 261
DB 870 GTG 872

RESULT 9
AAF92116
ID AAF92116 standard; cDNA; 2121 BP.
XX
AC AAF92116;
XX
DT 15-MAY-2001 (first entry)
XX
DE Human PRO1572 cDNA.
XX
KW Human; PRO protein; mapping; ss.
XX
OS Homo sapiens.
XX
PN WO200116318-A2.
XX
PD 08-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-US023328.
XX
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 07-DEC-1999; 99US-0169495P.
PR 09-DEC-1999; 99US-0170262P.
PR 11-JAN-2000; 2000US-0175481P.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 01-MAR-2000; 2000WO-US005601.
PR 03-MAR-2000; 2000US-0187202P.
PR 21-MAR-2000; 2000US-0191007P.
PR 30-MAR-2000; 2000WO-US008439.
PR 25-APR-2000; 2000US-0199397P.
PR 22-MAY-2000; 2000WO-US014042.
PR 05-JUN-2000; 2000US-0209832P.
XX
PA (GETH) GENENTECH INC.
XX
XX Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2001-183260/18.
DR P-PSDB; AAB87584.
XX
PT Eighty four nucleic acids encoding PRO polypeptides, useful in molecular
PT biology, including use as hybridization probes, and in chromosome and
PT gene mapping.
XX
PS Claim 2; Fig 117; 278pp; English.
XX

PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101916P.
PR 30-SEP-1998; 98US-0102570P.
PR 06-OCT-1998; 98US-0103449P.
PR 08-MAR-1999; 99WO-US005028.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021194.
PR 22-DEC-1999; 99WO-US030720.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 01-MAR-2000; 2000WO-US005601.
PR 30-MAR-2000; 2000WO-US008439.
PR 22-MAY-2000; 2000WO-US014042.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032378.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 30-MAY-2001; 2001WO-US017443.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.

(GETH) GENENTECH INC.

Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;

XX
DR WPI; 2002-731348/79.
DR P-PSDB; ABG95909.

XX
PT New isolated secreted and transmembrane PRO polypeptide useful for
PT modulating biological activity of a cell, or for treating sports-related
PT joint problems, osteoarthritis or rheumatoid arthritis.

XX
PS Claim 2; Fig 117; 399pp; English.

XX
CC The invention relates to an isolated secreted and transmembrane PRO
CC polypeptide having 80 % sequence identity to a sequence appearing as
CC ABG95851-ABG95934 or their associated signal peptide, or a sequence of an
CC extracellular domain of the proteins with their associated signal peptide
CC or lacking its associated signal peptide. Also included are the nucleic
CC acids encoding the proteins, vectors, host cells, fusion proteins and
CC antibodies which specifically bind to the proteins. The proteins are
CC useful for detecting a polypeptide designated as A, B, C or D in a sample
CC suspected of containing an A, B, C or D polypeptide, by contacting the
CC sample with a polypeptide designated as E, F, G, H or I (or vice versa)
CC and determining the formation of a A/E, B/F, B/G, C/H or D/I polypeptide
CC conjugate in the sample, where the formation of the conjugate is
CC indicative of the presence of an A, B, C or D polypeptide in the sample,
CC where A is a PRO10272 polypeptide, B is a PRO20110 polypeptide, C is a
CC PRO10096 polypeptide, D is a PRO19760 polypeptide, E is a PRO5801
CC polypeptide, F is a PRO1 polypeptide, G is a PRO20040 polypeptide, H is a
CC PRO20233 polypeptide and I is a PRO1890 polypeptide. The sample comprises
CC a cell suspected of expressing the A, B, C or D polypeptide. The E, F, G,
CC H or I polypeptide is labeled with a detectable label or is attached to a
CC solid support. The proteins are useful for linking a bioactive molecule
CC to a cell expressing a polypeptide designated as A, B, C or D or E, F, G,
CC H or I. The bioactive molecule is a toxin, a radiolabel or an antibody.
CC The bioactive molecule causes death of the cell. A, B, C, D, E, F, G, H,
CC or I, or antibodies against them are useful for modulating a biological
CC activity of a cell expressing a polypeptide designated as A, B, C or D or
CC E, F, G, H, or I. The cell is killed. The proteins are useful for
CC identifying agonists or antagonists, for the preparation of a medicament

CC useful in the treatment of a condition which is responsive to the
CC proteins, as molecular weight markers for protein electrophoresis
CC purposes, and as therapeutic agents for treating sports-related joint
CC problems, articular cartilage defects, osteoarthritis or rheumatoid
CC arthritis. Nucleic acids encoding the proteins are useful as
CC hybridisation probes, in chromosome and gene mapping, in the generation
CC of anti-sense RNA and DNA, for the preparation of the proteins, to
CC generate transgenic or knockout animals which are useful in the
CC development and screening of therapeutic useful reagents, for chromosome
CC identification, and in gene therapy. The antibody is useful as a
CC therapeutic agent, in a diagnostic assay and for affinity purification of
CC the protein from recombinant cell culture natural sources. The present
CC sequence encodes a novel secreted or transmembrane protein of the
CC invention

XX
SQ Sequence 2121 BP; 569 A; 516 C; 487 G; 549 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.05e-148 Length: 2121
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-063-731-118 (1-261) x ABS74436 (1-2121)

QY 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
90 ATGTCACACCACATGCCAAGTGGTGGCGTTCTCTCTGTCATCTTGGGGCTGGCGGC 149
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAsnProVal 40
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
150 TGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCGGTC 209
QY 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
210 ACCTCCGTGTTCCAGTACGAAGGCTCTGGAGGAGCTCGTGAGGCAGAGTTCAGGCTTC 269
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
270 ACCGAATGAGGCCCTATTTCACCATCTCTGGGACTTCCAGCCATGCTGCAGGCAGTGGCA 329
QY 81 AlaLeuMetIleValGlyIleValleuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
330 GCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCCTCTGGTATCCATCTTGCC 389
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
390 CTGAAATGCATCCGCATTTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACACTGACC 449
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
450 TCCGGGATCATGTTTCAATTGCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 509
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
510 AACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGGG 569
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
570 ATGGTGACAGCTGTTTCAGACCAGGTACACATTGGTGCGGCTCTGTTGGGCTGGGTC 629
QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
630 GCTGGAGGCCTCACACTAATTGGGGTGTGATGATGTGATCGCTGCCCTGCCGGGCGCTGGCA 689
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
690 CCAGAAGAAACCAACTACAAGCCGTTTCTTATCATGCTCTCAGGCCACAGTGTTCCTTAC 749
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||

| | | | |
|-----------|--------------|--|-----|
| Db | 750 | AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCTCAACCAACCAAAAAACAAGAAGATA | 809 |
| Qy | 241 | TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr | 260 |
| Db | 810 | TACGATGGAGGTGCCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCCAAGCAGCACTAT | 869 |
| Qy | 261 | Val 261 | |
| Db | 870 | GTG 872 | |
| RESULT 11 | | | |
| ABX78705 | | | |
| ID | ABX78705 | standard; cDNA; 2121 BP. | |
| XX | AC | ABX78705; | |
| XX | DT | 15-APR-2003 (first entry) | |
| XX | DE | Human PRO polynucleotide #178. | |
| KW | KW | Human; PRO; gene; ss; cytostatic; tumour; cancer; breast; lung; stomach; | |
| KW | KW | liver; dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT; | |
| KW | KW | antibody-dependent enzyme mediated prodrug therapy. | |
| XX | OS | Homo sapiens. | |
| XX | PN | US2003027272-A1. | |
| XX | PD | 06-FEB-2003. | |
| PF | 21-JUN-2002; | 2002US-00176492. | |
| XX | 18-SEP-1997; | 97US-0059263P. | |
| PR | 18-SEP-1997; | 97US-0059266P. | |
| PR | 17-OCT-1997; | 97US-0062250P. | |
| PR | 21-OCT-1997; | 97US-0063486P. | |
| PR | 24-OCT-1997; | 97US-0063120P. | |
| PR | 24-OCT-1997; | 97US-0063121P. | |
| PR | 28-OCT-1997; | 97US-0063540P. | |
| PR | 28-OCT-1997; | 97US-0063541P. | |
| PR | 28-OCT-1997; | 97US-0063544P. | |
| PR | 28-OCT-1997; | 97US-0063564P. | |
| PR | 29-OCT-1997; | 97US-0063734P. | |
| PR | 31-OCT-1997; | 97US-0063870P. | |
| PR | 31-OCT-1997; | 97US-0064103P. | |
| PR | 13-NOV-1997; | 97US-0065311P. | |
| PR | 21-NOV-1997; | 97US-0066120P. | |
| PR | 24-NOV-1997; | 97US-0066466P. | |
| PR | 11-DEC-1997; | 97US-0069335P. | |
| PR | 12-DEC-1997; | 97US-0069425P. | |
| PR | 17-DEC-1997; | 97US-0069870P. | |
| PR | 18-DEC-1997; | 97US-0068017P. | |
| PR | 10-MAR-1998; | 98US-0077450P. | |
| PR | 11-MAR-1998; | 98US-0077632P. | |
| PR | 11-MAR-1998; | 98US-0077649P. | |
| PR | 20-MAR-1998; | 98US-0078886P. | |
| PR | 20-MAR-1998; | 98US-0078939P. | |
| PR | 27-MAR-1998; | 98US-0079664P. | |
| PR | 31-MAR-1998; | 98US-0079786P. | |
| PR | 31-MAR-1998; | 98US-0080107P. | |
| PR | 31-MAR-1998; | 98US-0080194P. | |
| PR | 01-APR-1998; | 98US-0080327P. | |
| PR | 01-APR-1998; | 98US-0080333P. | |
| PR | 08-APR-1998; | 98US-0081049P. | |
| PR | 08-APR-1998; | 98US-0081070P. | |
| PR | 09-APR-1998; | 98US-0081195P. | |
| PR | 15-APR-1998; | 98US-0081838P. | |
| PR | 21-APR-1998; | 98US-0082568P. | |
| PR | 21-APR-1998; | 98US-0082569P. | |
| PR | 22-APR-1998; | 98US-0082704P. | |
| PR | 22-APR-1998; | 98US-0082797P. | |
| PR | 28-APR-1998; | 98US-0083322P. | |
| PR | 29-APR-1998; | 98US-0083495P. | |
| PR | 29-APR-1998; | 98US-0083496P. | |
| PR | 29-APR-1998; | 98US-0083499P. | |
| PR | 29-APR-1998; | 98US-0083559P. | |
| PR | 05-MAY-1998; | 98US-0084366P. | |
| PR | 06-MAY-1998; | 98US-0084414P. | |
| PR | 07-MAY-1998; | 98US-0084639P. | |
| PR | 07-MAY-1998; | 98US-0084640P. | |
| PR | 07-MAY-1998; | 98US-0084643P. | |
| PR | 15-MAY-1998; | 98US-0085579P. | |
| PR | 15-MAY-1998; | 98US-0085580P. | |
| PR | 15-MAY-1998; | 98US-0085582P. | |
| PR | 15-MAY-1998; | 98US-0085700P. | |
| PR | 18-MAY-1998; | 98US-0086023P. | |
| PR | 22-MAY-1998; | 98US-0086392P. | |
| PR | 22-MAY-1998; | 98US-0086486P. | |
| PR | 28-MAY-1998; | 98US-0087098P. | |
| PR | 28-MAY-1998; | 98US-0087208P. | |
| PR | 02-JUN-1998; | 98US-0087609P. | |
| PR | 02-JUN-1998; | 98US-0087759P. | |
| PR | 03-JUN-1998; | 98US-0087827P. | |
| PR | 04-JUN-1998; | 98US-0088025P. | |
| PR | 04-JUN-1998; | 98US-0088028P. | |
| PR | 04-JUN-1998; | 98US-0088029P. | |
| PR | 04-JUN-1998; | 98US-0088033P. | |
| PR | 04-JUN-1998; | 98US-0088326P. | |
| PR | 05-JUN-1998; | 98US-0088167P. | |
| PR | 05-JUN-1998; | 98US-0088202P. | |
| PR | 05-JUN-1998; | 98US-0088212P. | |
| PR | 05-JUN-1998; | 98US-0088217P. | |
| PR | 09-JUN-1998; | 98US-0088655P. | |
| PR | 10-JUN-1998; | 98US-0088722P. | |
| PR | 10-JUN-1998; | 98US-0088738P. | |
| PR | 10-JUN-1998; | 98US-0088740P. | |
| PR | 10-JUN-1998; | 98US-0088811 | |

PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Alignment Scores:

Pred. No.: 1.05e-148
Score: 1357.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%

Length: 2121
Matches: 261
Conservative: 0
Mismatches: 0
Indels: 0

DB: 7 Gaps: 0
US-10-063-731-118 (1-261) x ABX78705 (1-2121)
QY 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 90 ATGTCACCAACACATGCCAAGTGGTGGCTTCTCTCTCTCCATCTCTGGGCTGGCCGC 149
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 150 TGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCGTC 209
QY 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
Db 210 ACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGAGCTAAGGTGAGGAGTTCAGGCTTC 269
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 270 ACCGAATGCAGGCCCTATTTCACCATCTGGGACTTCCAGGCATGCTGCAGGCAGTGC 329
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 330 GCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTGGCCTCCTGGTATCCATCTTTGCC 389
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 390 CTGAATGCATCCGCATTGGCAGCATGGAGGACTCTGCAAGCCAAACATGACACTGACC 449
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 450 TCCGGGATCATGTTTCATTGTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 509
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 510 AACATGCTGGTGACTTAACCTTCGGATGTCCACAGCTAACATGTACACCGGCATGGTGG 569
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 570 ATGGTGACAGACTGTTTCAGACCAGGTACACATTGGTGGGCTCTGTTCTGGGCTGGGTC 629
QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 630 GCTGGAGGCTCACACTAATTGGGGGTGTGATGATGTGATCGCCTGCCGGGCTGGCA 689
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 690 CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTAC 749
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
Db 750 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAACAAGAGATA 809
QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 810 TACGATGGAGGTGCCCGCACAGAGACGAGGTACAATCTTATCTTATCTTCCAAGCAGCTAT 869
QY 261 Val 261
Db 870 GTG 872

RESULT 12

ACA75677
ID ACA75677 standard; cdna; 2121 BP.
XX
AC ACA75677;
XX
DT 07-JUL-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1572 cdna.
XX
KW Human; secreted and transmembrane protein: PRO; gene therapy;
KW tumour necrosis factor-alpha release; TNF-alpha release;
KW chondrocyte proliferation; chondrocyte differentiation; tumour;

PR 21-NOV-1997; 97US-00666120P.
PR 24-NOV-1997; 97US-00666466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-00693335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 16-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.

PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 24-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.

PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Alignment Scores:

Pred. No.: 1.05e-148 Length: 2121
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-10-063-731-118 (1-261) x ACA71157 (1-2121)

QY 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
DB 90 ATGTCCACCACACATGCCAAGTGGTGGGTTCTCTCTGTCCATCCTGGGCTGGCCGGC 149
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
DB 150 TGCATCGCGGCCACCGGATGGACATGGAGCACCAGGACCTGTACGACCAACCCCGTC 209
QY 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
DB 210 ACCTCCGTGTTCAGTACGAAGGGCTCTGGAGGAGCTGCGTAGGCAGAGTTCAGGCTTC 269
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
DB 270 ACCGAATGCAGGCCCTATTTCACCATCTGGGACTTCCAGCCATGCTGCAGGCAGTGGCA 329
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
DB 330 GCCCTGATGATCGTAGGCATCGTCTGGTGGCATTGGCCTCCTGGTATCCATCTTTGCC 389
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAasnMetThrLeuThr 120
DB 390 CTGAATGCATCCGATTTGGCAGCATGGAGGACTCTGCCAAGCCAACATGACACTGACC 449
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
DB 450 TCCGGGATCATGTTTCATTGTCTCAGGTTTGTGCAATTGCTGGAGTGTCTGTGTTTGGC 509
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAasnMetTyrThrGlyMetGlyGly 160
DB 510 AACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGGG 569
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
DB 570 ATGGTGCAGACTGTTCAGACCAGGTACACATTGGTGGGCTCTGTTGGGCTGGGTC 629
QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
DB 630 GCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCTGCCGGGGCCTGGCA 689
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220

Db 690 CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTTAC 749
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
Db 750 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAAACACCAAAACAAGAAGATA 809
QY 241 TyrAspGlyGlyAlaAArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 810 TACGATGGAGGTGCCCGCACAGAGGACGAGGTACATCTTATCTTCCAAGCAGCACTAT 869
QY 261 Val 261
Db 870 GTG 872
RESULT 14
ACC87685
ID ACC87685 standard; cDNA; 2121 BP.
XX
AC ACC87685;
XX
DT 05-AUG-2003 (first entry)
XX Human secreted polypeptide PRO1572-encoding cDNA, SEQ ID NO:355.
DE Human; PRO; secreted protein; transmembrane protein;
XX extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnery; gene therapy; gene; ss.
XX Homo sapiens.
OS
XX US2003027278-A1.
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176987.
XX 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.

PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 05-JUN-1998; 98US-0088655P.
PR 09-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.

PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Alignment Scores:
Pred. No.:

1.05e-148

Length:

2121

| | | | |
|---|--|---|-----|
| Score: 1357.00 | | Matches: 261 | |
| Percent Similarity: 100.00% | | Conservative: 0 | |
| Best Local Similarity: 100.00% | | Mismatches: 0 | |
| Query Match: 100.00% | | Indels: 0 | |
| DB: 7 | | Gaps: 0 | |
| US-10-063-731-118 (1-261) x ACC87685 (1-2121) | | | |
| QY | 1 | MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly | 20 |
| Db | 90 | ATGTCACCAACACATGCCAAGTGGTGGCGTTCTCTCTGTCCATCCTGGGGCTGGCCGGC | 149 |
| QY | 21 | CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal | 40 |
| Db | 150 | TGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTC | 209 |
| QY | 41 | ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe | 60 |
| Db | 210 | ACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGGGTGAGGCAGAGTTCAGGCTTC | 269 |
| QY | 61 | ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg | 80 |
| Db | 270 | ACCGAATGCAGGCCCTATTTCACCATCTGGGACTTCAGGCCATGCTGCAGGCACTGCCGA | 329 |
| QY | 81 | AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla | 100 |
| Db | 330 | GCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTGGGCTCCTGGTATCCATCTTTGCC | 389 |
| QY | 101 | LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr | 120 |
| Db | 390 | CTGAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACTGACC | 449 |
| QY | 121 | SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla | 140 |
| Db | 450 | TCCGGGATCATGTTTCATTGTCTCAGGCTTTGTGCAATTGCTGGAGTGTCTGTGTTGCC | 509 |
| QY | 141 | AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly | 160 |
| Db | 510 | AACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG | 569 |
| QY | 161 | MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal | 180 |
| Db | 570 | ATGGTGACAGACTGTTCAGACCAGGTACACATTGGTGCGGCTCTGTCGTGGGCTGGGTC | 629 |
| QY | 181 | AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla | 200 |
| Db | 630 | GCTGGAGGCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGCTGGCA | 689 |
| QY | 201 | ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr | 220 |
| Db | 690 | CCAGAAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTTCCTAC | 749 |
| QY | 221 | LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle | 240 |
| Db | 750 | AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAAACACCAAAACAAAGAAGATA | 809 |
| QY | 241 | TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr | 260 |
| Db | 810 | TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCTTCCAAGCACCACTAT | 869 |
| QY | 261 | Val 261 | |
| Db | 870 | GTG 872 | |
| RESULT 15 | | | |
| ACC87071 | | | |
| ID | ACC87071 standard; cDNA; 2121 BP. | | |
| XX | | | |
| AC | ACC87071; | | |
| XX | | | |
| DT | 05-AUG-2003 (first entry) | | |
| XX | | | |
| DE | Human secreted polypeptide PRO1572-encoding cDNA, SEQ ID NO:355. | | |

| | |
|----|--|
| XX | Human; PRO; secreted protein; transmembrane protein; TNF-alpha; extracellular domain; tumour necrosis factor-alpha; cartilage disorder; chondrocyte; proliferation; differentiation; cancer; tumour; diagnosis; bone disorder; arthritis; sports injury; cancer; tumour; diagnosis; adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix; liver; drug screening; transgenic animal; genetic analysis; antiarthritic; vulnery; gene therapy; gene; ss. |
| XX | Homo sapiens. |
| OS | US2003036159-A1. |
| XX | 20-FEB-2003. |
| PN | 02-JUL-2002; 2002US-00188773. |
| XX | 18-SEP-1997; 97US-0059263P. |
| PR | 18-SEP-1997; 97US-0059266P. |
| PR | 17-OCT-1997; 97US-0062250P. |
| PR | 21-OCT-1997; 97US-0063486P. |
| PR | 24-OCT-1997; 97US-0063120P. |
| PR | 24-OCT-1997; 97US-0063121P. |
| PR | 28-OCT-1997; 97US-0063540P. |
| PR | 28-OCT-1997; 97US-0063541P. |
| PR | 28-OCT-1997; 97US-0063544P. |
| PR | 28-OCT-1997; 97US-0063564P. |
| PR | 29-OCT-1997; 97US-0063734P. |
| PR | 31-OCT-1997; 97US-0063870P. |
| PR | 31-OCT-1997; 97US-0064103P. |
| PR | 13-NOV-1997; 97US-0065311P. |
| PR | 21-NOV-1997; 97US-0066120P. |
| PR | 24-NOV-1997; 97US-0066466P. |
| PR | 24-NOV-1997; 97US-0066772P. |
| PR | 11-DEC-1997; 97US-0069335P. |
| PR | 12-DEC-1997; 97US-0069425P. |
| PR | 17-DEC-1997; 97US-0069870P. |
| PR | 18-DEC-1997; 97US-0068017P. |
| PR | 10-MAR-1998; 98US-0077450P. |
| PR | 11-MAR-1998; 98US-0077632P. |
| PR | 11-MAR-1998; 98US-0077649P. |
| PR | 20-MAR-1998; 98US-0078886P. |
| PR | 20-MAR-1998; 98US-0078939P. |
| PR | 27-MAR-1998; 98US-0079664P. |
| PR | 27-MAR-1998; 98US-0079786P. |
| PR | 31-MAR-1998; 98US-0080107P. |
| PR | 31-MAR-1998; 98US-0080194P. |
| PR | 01-APR-1998; 98US-0080327P. |
| PR | 01-APR-1998; 98US-0080333P. |
| PR | 08-APR-1998; 98US-0081049P. |
| PR | 08-APR-1998; 98US-0081070P. |
| PR | 09-APR-1998; 98US-0081195P. |
| PR | 15-APR-1998; 98US-0081838P. |
| PR | 21-APR-1998; 98US-0082568P. |
| PR | 21-APR-1998; 98US-0082569P. |
| PR | 22-APR-1998; 98US-0082704P. |
| PR | 22-APR-1998; 98US-0082797P. |
| PR | 28-APR-1998; 98US-0083322P. |
| PR | 29-APR-1998; 98US-0083495P. |
| PR | 29-APR-1998; 98US-0083496P. |
| PR | 29-APR-1998; 98US-0083499P. |
| PR | 29-APR-1998; 98US-0083559P. |
| PR | 05-MAY-1998; 98US-0084366P. |
| PR | 06-MAY-1998; 98US-0084414P. |
| PR | 07-MAY-1998; 98US-0084639P. |
| PR | 07-MAY-1998; 98US-0084640P. |
| PR | 07-MAY-1998; 98US-0084643P. |
| PR | 15-MAY-1998; 98US-0085579P. |
| PR | 15-MAY-1998; 98US-0085580P. |
| PR | 15-MAY-1998; 98US-0085582P. |
| PR | 15-MAY-1998; 98US-0085700P. |
| PR | 18-MAY-1998; 98US-0086023P. |
| PR | 22-MAY-1998; 98US-0086392P. |

PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 26-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.

PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 1.05e-148 | Length: | 2121 |
| Score: | 1357.00 | Matches: | 261 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 7 | Gaps: | 0 |

US-10-063-731-118 (1-261) x ACC87071 (1-2121)

| | | | |
|----|-----|--|-----|
| QY | 1 | MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly | 20 |
| Db | 90 | ATGTCCACCAACACATGCCAAGTGGTGGCTTCTCTCTCTCCATCTTGGGGCTGGCCGGC | 149 |
| QY | 21 | CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal | 40 |
| Db | 150 | TGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTC | 209 |
| QY | 41 | ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe | 60 |
| Db | 210 | ACCTCCGTGTTCAGTACGAAGGGCTCTGGAGGAGCTCGGTGAGGCAGAGTTCAGGCTTC | 269 |
| QY | 61 | ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg | 80 |
| Db | 270 | ACCGAATGCAGGCCCTATTTCACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCGA | 329 |

QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 330 GCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTGGCCCTCTGGTATCCATCTTTGCC 389

QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 390 CTGAATGCAATCCGCATTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACACTGACC 449

QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 450 TCCGGGATCATGTTTCATTGTTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTGCC 509

QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 510 AACATGCTGTGTACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGGG 569

QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 570 ATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGGCTCTGTTCTGTTGGGCTGGGTC 629

QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 630 GCTGGAGGCTCACACTAATTGGGGGTGTGATGATGTGCATGCGCTGCCGGGGCCCTGGCA 689

QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 690 CCAGAAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTTCCTAC 749

QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
Db 750 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAAACACCAAAACAAAGAGATA 809

QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 810 TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCTTCAAGCACGACTAT 869

QY 261 Val 261
Db 870 GTG 872

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 14:25:29 ; Search time 3730 Seconds
(without alignments)
16980.613 Million cell updates/sec

Title: US-10-063-731-117
Perfect score: 2121
Sequence: 1 gagctccctcaggagcg.....ggtaaaactaattcttaa 2121

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 786 | 37.1 | 786 | 29 | AY402516 |
| 2 | 785.2 | 37.0 | 1751 | 11 | AK033657 |
| 3 | 694.2 | 32.7 | 975 | 13 | BU838449 |
| 4 | 674.2 | 31.8 | 679 | 12 | BM791434 |

| | | | | | |
|----|-------|------|------|----|----------|
| 5 | 664 | 31.3 | 675 | 12 | BG547464 |
| 6 | 656.8 | 31.0 | 743 | 12 | BM989992 |
| 7 | 647.2 | 30.5 | 1138 | 14 | CF110345 |
| 8 | 623 | 29.4 | 680 | 12 | BM759274 |
| 9 | 607.6 | 28.6 | 795 | 29 | AY402518 |
| 10 | 593.2 | 28.0 | 617 | 29 | AY402517 |
| 11 | 588.2 | 27.7 | 627 | 12 | BQ016897 |
| 12 | 552 | 26.0 | 743 | 12 | BI823718 |
| 13 | 525.8 | 24.8 | 681 | 14 | CB462343 |
| 14 | 519 | 24.5 | 651 | 12 | BM837765 |
| 15 | 495.8 | 23.4 | 653 | 12 | BM788945 |
| 16 | 434 | 20.5 | 549 | 12 | BM791457 |
| 17 | 413.2 | 19.5 | 961 | 12 | BG538299 |
| 18 | 394 | 18.6 | 567 | 13 | BX528145 |
| 19 | 386 | 18.2 | 627 | 10 | BB663682 |
| 20 | 378 | 17.8 | 540 | 12 | BM836731 |
| 21 | 377.2 | 17.8 | 600 | 12 | BM997852 |
| 22 | 374.6 | 17.7 | 797 | 13 | BU946638 |
| 23 | 373.6 | 17.6 | 421 | 10 | BF091029 |
| 24 | 360.8 | 17.0 | 379 | 10 | BE841291 |
| 25 | 349.6 | 16.5 | 420 | 14 | N99013 |
| 26 | 346.8 | 16.4 | 458 | 10 | BE841654 |
| 27 | 341 | 16.1 | 341 | 12 | BM836109 |
| 28 | 340.2 | 16.0 | 412 | 10 | AW867300 |
| 29 | 339.6 | 16.0 | 402 | 12 | BG937956 |
| 30 | 339 | 16.0 | 493 | 14 | CA436277 |
| 31 | 339 | 16.0 | 552 | 12 | BM819720 |
| 32 | 339 | 16.0 | 567 | 12 | BM819706 |
| 33 | 338.8 | 16.0 | 570 | 12 | BM794650 |
| 34 | 337.4 | 15.9 | 458 | 13 | BX108271 |
| 35 | 336 | 15.8 | 395 | 9 | AA961052 |
| 36 | 335.8 | 15.8 | 585 | 12 | BM785703 |
| 37 | 335.8 | 15.8 | 661 | 12 | BM791792 |
| 38 | 335.6 | 15.8 | 633 | 10 | BB625263 |
| 39 | 335.4 | 15.8 | 403 | 9 | AI926840 |
| 40 | 334.4 | 15.8 | 483 | 9 | AI187329 |
| 41 | 334.2 | 15.8 | 593 | 12 | BM829469 |
| 42 | 333.8 | 15.7 | 367 | 9 | AW027675 |
| 43 | 315.4 | 14.9 | 649 | 10 | BB529462 |
| 44 | 308.2 | 14.5 | 548 | 9 | AA882221 |
| 45 | 308 | 14.5 | 731 | 29 | CE210717 |
| 46 | 305.8 | 14.4 | 660 | 29 | AG148436 |
| 47 | 301 | 14.2 | 622 | 10 | BF426746 |
| 48 | 300.4 | 14.2 | 428 | 10 | AW319069 |
| 49 | 295 | 13.9 | 484 | 9 | AA028634 |
| 50 | 287.2 | 13.5 | 500 | 14 | CD244706 |
| 51 | 283.8 | 13.4 | 441 | 12 | BM839206 |
| 52 | 268.2 | 12.6 | 633 | 10 | BF426867 |
| 53 | 249.2 | 11.7 | 374 | 13 | BY063182 |
| 54 | 248.8 | 11.7 | 1069 | 14 | CB596764 |
| 55 | 236.8 | 11.2 | 455 | 12 | BM751562 |
| 56 | 231 | 10.9 | 232 | 14 | N99040 |
| 57 | 219.4 | 10.3 | 262 | 9 | AI872593 |
| 58 | 202 | 9.5 | 520 | 14 | CB220415 |
| 59 | 188.2 | 8.9 | 550 | 13 | BQ210042 |
| 60 | 183 | 8.6 | 642 | 28 | AQ426747 |
| 61 | 183 | 8.6 | 740 | 14 | CB597222 |
| 62 | 182.2 | 8.6 | 3868 | 11 | BC033908 |
| 63 | 179.6 | 8.5 | 670 | 29 | AG086778 |
| 64 | 179 | 8.4 | 449 | 9 | AI365624 |
| 65 | 178.2 | 8.4 | 704 | 28 | BZ598866 |
| 66 | 177.6 | 8.4 | 884 | 13 | BU187437 |
| 67 | 177 | 8.3 | 660 | 29 | AG142871 |
| 68 | 176.4 | 8.3 | 724 | 29 | AG117953 |
| 69 | 176 | 8.3 | 424 | 14 | H01809 |
| 70 | 176 | 8.3 | 461 | 10 | AW817385 |
| 71 | 175.8 | 8.3 | 978 | 12 | BM475038 |
| 72 | 175.6 | 8.3 | 333 | 9 | AA730646 |
| 73 | 174.6 | 8.2 | 466 | 14 | H07128 |
| 74 | 174.2 | 8.2 | 544 | 13 | BX491755 |
| 75 | 174 | 8.2 | 774 | 28 | AQ745937 |
| 76 | 173.4 | 8.2 | 662 | 29 | AG182637 |
| 77 | 173.2 | 8.2 | 649 | 13 | BU617622 |

78 173.2 8.2 819 29 CG734227 RP11-224F
79 172.8 8.1 664 28 AQ491423 PCI-11-2
80 172.6 8.1 358 28 AQ665179 HS_5344_B
81 172.4 8.1 185 10 BE937032 PM3-SN002
82 172.2 8.1 654 29 AG042723 Pan trogl
83 172.2 8.1 696 13 BX485530 DKFZp686I
84 172 8.1 685 29 AG187749 Pan trogl
85 171.8 8.1 630 28 AQ346241 RPCI11-11
86 171.2 8.1 700 29 AG114100 Pan trogl
87 171.2 8.1 711 9 AV376479
88 171.2 8.1 720 10 BF037410
89 171 8.1 527 13 BU740001 UI-E-EO0-
90 170.2 8.0 445 28 AQ664011 HS_2277_B
91 170 8.0 629 28 AQ422183 PCI-11-2
92 170 8.0 745 29 AG015273 Homo sapi
93 170 8.0 750 29 AG015272 Homo sapi
94 170 8.0 810 28 AQ780966 HS_3138_B
95 169.8 8.0 398 14 H75272 YU07901.F1
96 169.6 8.0 422 28 B82626
97 169.6 8.0 486 28 AQ546155 CITBI-B1-
98 169.6 8.0 682 10 BF677863
99 169.4 8.0 441 9 AA859001 UI-R-A0-b
100 169.4 8.0 533 28 AQ626520 CITBI-B1-

ALIGNMENTS

RESULT 1
AY402516 786 bp DNA linear GSS 15-DEC-2003
LOCUS Homo sapiens CLDN18 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402516
VERSION AY402516.1 GI:39758502
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 786)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene tricos
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 786)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>786
/gene="CLDN18"
/locus_tag="HCM1250"
ORIGIN
Query Match 37.1%; Score 786; DB 29; Length 786;
Best Local Similarity 100.0%; Pred. No. 6.1e-153;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 ATGTCCACCACCATGCAAGTGGTGGCGTTCTCTGTCTCCATCCTGGSGCTGGCCGGC 149
|||||
Db 1 ATGTCCACCACCATGCAAGTGGTGGCGTTCTCTGTCTCCATCCTGGSGCTGGCCGGC 60
|||||
QY 150 TGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAAACCCCGTC 209
|||||
Db 61 TGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAAACCCCGTC 120
|||||
QY 210 ACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTCGTGAGGCAGAGTTCAGGCTTC 269
|||||
Db 121 ACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTCGTGAGGCAGAGTTCAGGCTTC 180
|||||
QY 270 ACCGAATGCAGGCCCTATTTCACCATCTCTGGGACTTCCAGCCATGTCAGGCAGTGC 329
|||||
Db 181 ACCGAATGCAGGCCCTATTTCACCATCTCTGGGACTTCCAGCCATGTCAGGCAGTGC 240
|||||
QY 330 GCCTGATGATCGTAGGCATCGTCTGGGTGCCATTGGCCCTCCTGGTATCCATCTTTGCC 389
|||||
Db 241 GCCTGATGATCGTAGGCATCGTCTGGGTGCCATTGGCCCTCCTGGTATCCATCTTTGCC 300
|||||
QY 390 CTGAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACACTGACC 449
|||||
Db 301 CTGAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACACTGACC 360
|||||
QY 450 TCCGGGATCATGTTCAATTGTTCTCAGGTCTTTTGCAATTGCTGGAGTGTCTGTGTTGCC 509
|||||
Db 361 TCCGGGATCATGTTCAATTGTTCTCAGGTCTTTTGCAATTGCTGGAGTGTCTGTGTTGCC 420
|||||
QY 510 AACATGCTGGTGACTAATCTCTGGATGTCCACAGTAACATGTACACCGGATGGGTGGG 569
|||||
Db 421 AACATGCTGGTGACTAATCTCTGGATGTCCACAGTAACATGTACACCGGATGGGTGGG 480
|||||
QY 570 ATGCTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGGCTCTGTTCTGGGCTGGGTC 629
|||||
Db 481 ATGCTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGGCTCTGTTCTGGGCTGGGTC 540
|||||
QY 630 GCTGGAGGCCTCACACTAATTGGGGTGTGATGATGTGCATCGCCTGCCGGGCTGGCA 689
|||||
Db 541 GCTGGAGGCCTCACACTAATTGGGGTGTGATGATGTGCATCGCCTGCCGGGCTGGCA 600
|||||
QY 690 CCAGAAAGAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTGCCTAC 749
|||||
Db 601 CCAGAAAGAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTGCCTAC 660
|||||
QY 750 AAGCCTGGAGGCTTCAAGCCAGCACTGGCTTTGGTCCCAACACCAAAACAAGAAGATA 809
|||||
Db 661 AAGCCTGGAGGCTTCAAGCCAGCACTGGCTTTGGTCCCAACACCAAAACAAGAAGATA 720
|||||
QY 810 TACGATGGAGGTGCCCCGACAGAGGACGAGGTACATCTTATCTTCCAGCAGCACTAT 869
|||||
Db 721 TACGATGGAGGTGCCCCGACAGAGGACGAGGTACATCTTATCTTCCAGCAGCACTAT 780
|||||
QY 870 GTGTAA 875
|||||
Db 781 GTGTAA 786
|||||
RESULT 2
AK033657
LOCUS Mus musculus adult male cecum cDNA, RIKEN full-length enriched
DEFINITION library, clone:9130209K06 product:CLAUDIN-18A2.1, full insert
sequence.
ACCESSION AK033657
VERSION AK033657.1 GI:26329346
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsumai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1751)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,I., Kasukawa,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kouda,M., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. .1751
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9130209K06"
/db_xref="MGI:2397735"
/db_xref="taxon:10090"
/clone="9130209K06"

| | | | |
|----|------|---|------|
| Db | 771 | AACAAGAAGATCTACGATGGGGTGTCCCGCACAGAAGACGATGAACAGTCTCATCTCTACC | 830 |
| Qy | 858 | AAGCAGCACTATGTGTAATGCTCTAAGACCTCTCAGCACGGCGGAAGAACTCCC---- | 913 |
| Db | 831 | AAGTATGACTATGTGTAGTCTCTAAGACCCGCGCAACCTGTGTGCAGGAGAACCTTCC | 890 |
| Qy | 914 | -----GGAGAGCTCACCCAAAAAACAAGGAGATCCCATCTAGATTTCTTCTGCTTTTGAC | 969 |
| Db | 891 | CCAAGAAGAGCTCACCC-CAAAAGCAACGGGAGTCTACCTTGTTCCTTGTGATTTCAAC | 949 |
| Qy | 970 | TCACAGCTGGAAGTTAGAAAAAGCCTCGATTTCTATCTTTGGAGAGCGCCAAATGGTCTTAGC | 1029 |
| Db | 950 | TGACATCTGAAAGTTGGTAAAGCCTGATTTTCATCCATAGGAGGCTAGACAGTCTTGGC | 1009 |
| Qy | 1030 | CTCA-GTCTCTGTCTCTAAATATTCACCAATAAACAGCTGAGTTA---TTTATGAATTA | 1085 |
| Db | 1010 | CACATGTGTCTGCCTCTAAATATCCCATCACAAAAAGCTGAGTTATCGTTTATGAGTTA | 1069 |
| Qy | 1086 | GAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTAAATATAACT--TTCTACT | 1143 |
| Db | 1070 | GAGGCCATAACACTCACTTTAGCCCAACCCCTCTGCTTTTACCGTAGACTTCTTTTCAT | 1129 |
| Qy | 1144 | CTGATGAGAGAATGTGGTTTAAATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCC | 1203 |
| Db | 1130 | CTGGTGATGGAATGGAATTTGACTCACAGACTAATCTTAATGGTTTAGAGAACTTTC | 1189 |
| Qy | 1204 | CCTCTTCTCTAGTCAATAAAACCCATTGATGATCTATTTCCCAGCTT-ATCCCCAAGAA | 1262 |
| Db | 1190 | ---CTTCTCTGTACTTAATAAGCCTGCTGATGGTTCGATTTTCCAGCTTGACCAACCAAGG | 1246 |
| Qy | 1263 | AACCTTTTGAAGGAAGAGTAGACCCCAAGATGTTATTTCTGCTGTTTGAATTTGTCT | 1322 |
| Db | 1247 | AAATTTTAAAGGAAAAAATAACATTAAGGCAATATTCTTACTCAATTTGTGCTT | 1306 |
| Qy | 1323 | CCCCACCCCACTTTGGC-----TAGTAATAACACTTACTGAAGAAGAGCAATAAGA | 1376 |
| Db | 1307 | ACCCACCCCAACTTGACTGATAATAATAATGAACACCACTTAAAGAAAGATGCCAGAG | 1366 |
| Qy | 1377 | GAAAGATATTGT-AATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATCTTAA | 1435 |
| Db | 1367 | GAAAGATAGTTGTGTTTCCCCCAGCCAGTCACTGAGTCCCCCTATGTGGTGATCTAGA | 1426 |
| Qy | 1436 | AAGTTACCAAAACCAAGTCAATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTGTTGA | 1495 |
| Db | 1427 | ACATTA-CTCGCCACAGTGATTTTCA-AAGAAGCAAGCGAGCCTGTTTCTGCTCAG | 1484 |
| Qy | 1496 | CATCTTCTTATTACAGCAACACCATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTC | 1555 |
| Db | 1485 | CATCTGCTGATTCCAGCAAGGCCCTTCCAGAGCTTTCCACTA-----GAAGTCTCTC | 1535 |
| Qy | 1556 | TTTCTGTCGGGTACAGAAATTTGCCCTAGATGAATGAGAAA---ATTATTTTTTTTAAT | 1612 |
| Db | 1536 | CTTCTCTCGGAAGTCAGAAAATTTCCCCCTAGAAAGAGTAAGAAATAGATTCTTTTGGGTAAC | 1595 |
| Qy | 1613 | TTAAGTCTTAAATATAGTTAAAAATAAATAATGTTTGTAGTAAATGATACACTATCTGT | 1672 |
| Db | 1596 | CTGAGTCTTAGGTATAGTTATAATAATAGTATATTAGCAAAACGGTTTGGTATCTCAGT | 1655 |
| Qy | 1673 | GAAATAGCCTCACCCCTACATGTGGATAGAAGGAATGAAAAATAATTTGCTTTTGACATT | 1732 |
| Db | 1656 | GAATTAGTTTCAGCCTTACATATAGAAAAAGCTGGGAAAAAAGCATCCCTTTGACATT | 1715 |
| Qy | 1733 | GTCTATATGGTA | 1744 |
| Db | 1716 | GTCTATAGCGTA | 1727 |

RESULT 3
BU838449
LOCUS
DEFINITION
AGENCOURT_8123082_lupski_dorsal_root_ganglion_Homo_sapiens_CDNA
clone IMAGE:6177081_5', mRNA sequence.
ACCESSION
BU838449

```

VERSION BU838449.1 GI:24022844
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 975)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13554 row: h column: 10
High quality sequence stop: 636.
Location/Qualifiers
1. .975
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6177081"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski_dorsal_root_ganglion"
/notes="vector: pCMV-SPORT6 (Life Technologies); Site_1:
Note1; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCTCCG-3' and
5'-GACTAGTTCTAGATCGGAGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
FEATURES
SOURCE

```

ORIGIN

| | Query Match | 32.7%; | Score 694.2; | DB 13; | Length 975; |
|----|-----------------------|--|--|-------------|-------------|
| | Best Local Similarity | 86.4%; | Pred. No. 7.5e-134; | | |
| | Matches 852; | Conservative 0; | Mismatches 14; | Indels 120; | Gaps 3; |
| QY | 81 | GCCAGGATCATGTCCACCA | CCACATGCCAAGTGGTGGCGTTCTCTCTGTCTCCATCTCTGGG | 140 | |
| DB | 1 | GCCAGGATCATGTCCACCA | CCACATGCCAAGTGGTGGCGTTCTCTCTGTCTCCATCTCTGGG | 60 | |
| QY | 141 | CTGGCCGGCTGCATCGCGGCC | ACCAGGATGGACATGTGGAGCACCCAGGACCTGTACGAC | 200 | |
| DB | 61 | CTGGCCGGCTGCATCGCGGCC | ACCAGGATGGACATGTGGAGCACCCAGGACCTGTACGAC | 120 | |
| QY | 201 | AACCCCGTCACTCCGTGTTCCAGTACGAAGGCTCTGGAGGAGTGCCTGAGGCAGAGT | 260 | | |
| DB | 121 | AACCCCGTCACTCCGTGTTCCAGTACGAAGGCTCTGGAGGAGTGCCTGAGGCAGAGT | 180 | | |
| QY | 261 | TCAGGCTTCACCGAATG | CAGGCCCTATTTCACCATCTCTGGGACTTCCAGCCATGCTGCAG | 320 | |
| DB | 181 | TCAGGCTTCACCGAATG | CAGGCCCTATTTCACCATCTCTGGGACTTCCAGCCATGCTGCAG | 240 | |
| QY | 321 | GCAGTGGAGCCCCTGATGATCGTAGGCATCGTCTCTGGGTGCCATTGGCCCTCCTGGTATCC | 380 | | |
| DB | 241 | GCAGTGGAGCCCCTGATGATCGTAGGCATCGTCTCTGGGTGCCATTGGCCCTCCTGGTATCC | 300 | | |
| QY | 381 | ATCTTTGCCCTGAAATGCATCCGCATTTGGCAGCATGGAGGACTCTGCCAAAGCCCAACATG | 440 | | |
| DB | 301 | ATCTTTGCCCTGAAATGCATCCGCATTTGGCAGCATGGAGGACTCTGCCAAAGCCCAACATG | 360 | | |
| QY | 441 | ACACTGACCTCCGGGATCATGTTTCATTTGCTCAGGTCTTTGTGCAATTGCTGGAGTGCT | 500 | | |

```
|||||
361 ACACAGCTCCGGGATCATGTTTCATGTCCT----- 391
QY 501 GTGTTTGCCAAACATGCTGGTGACCTAACTTCTGGATGTCCACAGCTAACATGTACACCGGC 560
Db 392 ----- 391
QY 561 ATGGGTGGGATGGTGACAGACTGTTCCAGACCAGGTACACATTTGGTGGCGGCTCTGTTGCTG 620
Db 392 -----CAGGTACACATTTGGTGGCGGCTCTGTTGCTG 422
QY 621 GGCTGGGTGCTGGAGGCCTCACACTAATTTGGGGTGTGATGATGTGCATCGCCTGCGGG 680
Db 423 GGCTGGGTGCTGGAGGCCTCACACTAATTTGGGGTGTGATGATGTGCATCGCCTGCGGG 482
QY 681 GGCCTGGCACCAGAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGT 740
Db 483 GGCCTGGCACCAGAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCGGGCCACAGT 542
QY 741 GTTGCTTACAGCCTGGAGGCTTCAAGCCAGCACTGGCTTTGGGTCCAAACACCAAAAAC 800
Db 543 GTTGCTTACAGCCTGGAGGCTTCAAGCCAGCACTGGCTTTGGGTCCAAACACCAAAAAC 602
QY 801 AAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAG 860
Db 603 AAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCCTT-CAAG 661
QY 861 CACGACTATGTCTAATGCTCTAAGACCTCTCAGCACGGCGGGAAGAAAACCTCCGGGAGAGC 920
Db 662 CACGACTATGTCTAATGCTCTAAGACCTCTCAGCACGGCGGGAAGAAAACCTCCGGGAGAGC 721
QY 921 TCACCCAAAAACAAGGAGATCCCATCTAGATTTT-CTTCTTGCTTTTGACTCAGAGCTGG 979
Db 722 TCACCCAAAAACAAGGAGATCCCATCCAGATTTCTTCTTGCTTNTGACTCAGAGCTGG 781
QY 980 AAGTTAGAAAAGCTCGATTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCT 1039
Db 782 AAGTTAGAAAAGCTCGATTCCTTCTTTGGAGAGGCCAAATGGGCTAGCCTCAGTCTCT 841
QY 1040 GTCTCTAAATATTCACCACATAAAACA 1065
Db 842 GTCTCTAAATATTCGCCCACTAAAAA 867
```

```
RESULT 4
BM791434
LOCUS
DEFINITION BM791434 679 bp mRNA linear EST 05-MAR-2002
K-EST0071443 S21SNU520 Homo sapiens cDNA clone S21SNU520-20-A11 5',
mRNA sequence.
ACCESSION BM791434
VERSION BM791434.1 GI:19139666
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 679)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.B., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 20 row: A column: 11
High quality sequence stop: 679.
FEATURES
source Location/Qualifiers
1..679
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNU520-20-A11"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNU520"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
```

ORIGIN

```
Query Match 31.8%; Score 674.2; DB 12; Length 679;
Best Local Similarity 99.6%; Pred. No. 1.1e-129;
Matches 676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 952 TTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTCATCTTTGGAG 1011
Db 1 TTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTCATCTTTGGAG 60

QY 1012 AGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACAGCTGAG 1071
Db 61 AGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACAGCTGAG 120

QY 1072 TTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCCTCAATTTCTTTTAAATA 1131
Db 121 TTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCCTCAATTTCTTTTAAATA 180

QY 1132 TAACTTTCTACTCTGATGAGAGAAATGTGGTTTAAATCTCTCTCACATTTTGTGATGATT 1191
Db 181 TAACTTTCTACTCTGATGAGAGAAATGTGGTTTAAATCTCTCTCTCACATTTTGTGATGATT 240

QY 1192 AGACAGACTCCCCCTCTTCTCCTAGTCAATAAACCCCATTTGATGATCTATTTCCCAGCTT 1251
Db 241 AGACAGACTCCCCCTCTTCTCCTAGTCAATAAACCCCATTTGATGATCTATTTCCCAGCTT 300

QY 1252 ATCCCCAAGAAAACCTTTTGAAGGAAGAGTAGACCCCAAGATGTTATTTTCTGCTGTTT 1311
Db 301 ATCCCCAAGAAAACCTTTTGAAGGAAGAGTAGACCCCAAGATGTTATTTTCTGCTGTTT 360

QY 1312 GAAATTTGTCTCCCAACCCCAACTTTGGCTAGTAATAAACACTTACTGAAAGAAGCA 1371
Db 361 GAAATTTGTCTCCCAACCCCAACTTTGGCTAGTAATAAACACTTACTGAAAGAAGCA 420

QY 1372 TAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATC 1431
Db 421 TAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATC 480

QY 1432 TTAAAAGTTACCAACCAAGTCATTTTCAGTTTGGGCAACCAAAACCTTTTCTACTGCTG 1491
Db 481 TTAAAAGTTACCAACCAAGTCATTTTTCAGTTTGGGCAACCAAAACCTTTTCTACTGCTG 540

QY 1492 TTGACATCTTCTTATTACAGCAACACCAATCTTAGGAGTTTCTCTGAGCTCTCCACTGGAGT 1551
Db 541 TTGACATCTTCTTATTACAGCAACACCAATCTTAGGAGTTTCTCTGAGCTCTCCACTGGAGT 600

QY 1552 CCTCTTCTGTGCGGGTCAAGAAATTGTCCCTAGATGAATGAGAAATATTTTAAATA 1611
```

```
|||||
601 CCTCTTCTGTACGGGTGAGAAATGTCCTAGATGAATGAGAAATATTATTTTAA 660
|||||
1612 TTAAAGTCCTAAATATAGT 1630
|||||
661 TTAAAGTCCTAAATATAGT 679

BG547464
LOCUS
DEFINITION
602574890F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4702823 5',
mRNA sequence.
ACCESSION
BG547464
VERSION
BG547464.1 GI:13546129
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1540 row: d column: 24
High quality sequence stop: 675.
Location/Qualifiers
1. 675
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4702823"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_77"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccgctcgcc); Site 2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 31.3%; Score 664; DB 12; Length 675;
Best Local Similarity 99.9%; Pred. No. 1.5e-127;
Matches 675; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 967 GACTCACAGCTGGAAGTTAGAAAAGCCTCGATTTCATCTTTGGAGAGGCCAAATGTCCT 1026
|||||
DB 1 GACTCACAGCTGGAAGTTAGAAAAGCCTCGATTTCATCTTTGGAGAGGCCAAATGTCCT 60
|||||
QY 1027 AGCCTCAGTCTCTGCTCTAAATATTCCACCATAAAACAGCTGAGTTATTATGAATTAG 1086
|||||
DB 61 AGCCTCAGTCTCTGCTCTAAATATTCCACCATAAAACAGCTGAGTTATTATGAATTAG 120
|||||
QY 1087 AGGCTATAGCTCACATTTTCAATCCTCTATTCTTTTAAATATAAATCTTCTACTCTG 1146
|||||
DB 121 AGGCTATAGCTCACATTTTCAATCCTCTATTCTTTTAAATATAAATCTTCTACTCTG 180
|||||
QY 1147 ATGAGAGAATGTTGTTTAAATCTCTCTCTACATTTTGATGATTTAGACAGACTCCCT 1206
|||||
```

```
Db 181 ATGAGAGAATGTGTTTAAATCTCTCTCTCACAATTTTGATGATTTAGACAGACTCCCT 240
QY 1207 CTTCTCTCTAGTCAATAAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACT 1266
|||||
DB 241 CTTCTCTCTAGTCAATAAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACT 300
|||||
QY 1267 TTTGAAAGGAAAGAGTAGACCCCAAGATGTTATTTTCTGCTGTTGTTGTTGTTCTCCCC 1326
DB 301 TTTGAAAGGAAAGAGTAGACCCCAAGATGTTATTTTCTGCTGTTGTTGTTGTTCTCCCC 360
|||||
QY 1327 ACCCCCAACTTGGCTAGTAATAAACACCTTACTGAAGAAGCAATAAAGAGAAATATT 1386
DB 361 ACCCCCAACTTGGCTAGTAATAAACACCTTACTGAAGAAGCAATAAAGAGAAATATT 420
|||||
QY 1387 TGTAATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATCTTAAAAAGTTACCAA 1446
DB 421 TGTAATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATCTTAAAAAGTTACCAA 480
|||||
QY 1447 CCAAAGTCATTTTTCAGTTTGAGGCAACCAAACTTTTCTACTGCTGTGACATCTTCTTAT 1506
DB 481 CCAAAGTCATTTTTCAGTTTGAGGCAACCAAACTTTTCTACTGCTGTGACATCTTCTTAT 540
|||||
QY 1507 TACAGCAACACCACTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTCTGTGCG 1566
DB 541 TACAGCAACACCACTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTCTGTGCG 600
|||||
QY 1567 GGTGAGAAATGTCCTAGATGAATGAGAAAAATTTTAAAAATTAAGTCTCTAAATA 1626
DB 601 GGTGAGAAATGTCCTAGATGAATGAGAAAAATTA-TTTTAAAAATTAAGTCTCTAAATA 659
|||||
QY 1627 TAGTTAAATAATAATA 1642
DB 660 TAGTTAAATAATAATA 675
|||||

RESULT 6
BM989992
LOCUS
DEFINITION
UI-H-D10-ato-h-11-0-UI.s1 NCI_CGAP_D10 Homo sapiens cDNA clone
IMAGE:5862202 3', mRNA sequence.
ACCESSION
BM989992
VERSION
BM989992.1 GI:19709381
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 743)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
sequence: 475-536, >AT-richLow_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. 743
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5862202"
/tissue_type="Lung Focal Fibrosis"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"

FEATURES
source
```

/clone lib="NCI_CGAP_D10"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI_CGAP_D10 is a cDNA library containing the following tissue(s): A pool of Lung Focal Fibrosis. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ATACGCGGTC.
TAG_TISSUE=lung with fibrosis
TAG_LIB=UI-H-D10
TAG_SEQ=ATACGCGGTC"

ORIGIN

Query Match 31.0%; Score 656.8; DB 12; Length 743;
Best Local Similarity 99.0%; Pred. No. 4.7e-126;
Matches 661; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1121 TTTTAAATATACTTTCTACTCTGATGAGAGAATGTGTTTAACTCTCTCTCACAT 1180
| | | | |
Db 1 TTTTCTTTTCTTTTCTACTCTGATGAGAGAATGTGTTTAACTCTCTCTCACAT 60
| | | | |

Qy 1181 TTTGATGATTTAGACAGACTCCCCCTCTTCTCCTAGTCAATAAACCCATTGATGATCTA 1240
| | | | |
Db 61 TTTGATGATTTAGACAGACTCCCCCTCTTCTCCTAGTCAATAAACCCATTGATGATCTA 120
| | | | |

Qy 1241 TTTCCAGCTTATCCCAAGAAAACCTTTTGAAAGAAAGAGTAGACCCAAAGATGTTATT 1300
| | | | |
Db 121 TTTCCAGCTTATCCCAAGAAAACCTTTTGAAAGAAAGAGTAGACCCAAAGATGTTATT 180
| | | | |

Qy 1301 TTCTGCTGTTGAATTTGTCTCCCAACCCCAACTTGGCTAGTAGTAATAAACACTTACTGA 1360
| | | | |
Db 181 TTCTGCTGTTGAATTTGTCTCCCAACCCCAACTTGGCTAGTAGTAATAAACACTTACTGA 240
| | | | |

Qy 1361 AGAAGAAGCAATAAGAGAAAGATATTGTAACTCTCCAGCCCATGATCTCGTTTCTTT 1420
| | | | |
Db 241 AGAAGAAGCAATAAGAGAAAGATATTGTAACTCTCCAGCCCATGATCTCGTTTCTTT 300
| | | | |

Qy 1421 ACACGTGATCTTAAAGTTACCAACCAAGTCATTTTCAGTTTGAGCAACCAACCT 1480
| | | | |
Db 301 ACACGTGATCTTAAAGTTACCAACCAAGTCATTTTCAGTTTGAGCAACCAACCT 360
| | | | |

Qy 1481 TTCTACTGCTGTTGACATCTTCTTATTACAGCAACACCATTTCTAGGAGTTTCTGAGCTC 1540
| | | | |
Db 361 TTCTACTGCTGTTGACATCTTCTTATTACAGCAACACCATTTCTAGGAGTTTCTGAGCTC 420
| | | | |

Qy 1541 TCCACTGGAGTCCTTTTCTGTGCGGGTCAAGAAATTGTCCCTAGATGAATGAGAAAATT 1600
| | | | |
Db 421 TCCACTGGAGTCCTTTTCTGTGCGGGTCAAGAAATTGTCCCTAGATGAATGAGAAAATT 480
| | | | |

Qy 1601 ATTTTAAATTTAAGTCCTAAATATAGTTAAATAAATAAATGTTTAAATGATA 1660
| | | | |
Db 481 ATTTTAAATTTAAGTCCTAAATATAGTTAAATAAATAAATGTTTAAATGATA 540
| | | | |

Qy 1661 CACTATCTCTGTGAATAGCCTCACCCCTACATGTGGATAGAAGGAATGAAAAATAAT 1720
| | | | |
Db 541 CACTATCTCTGTGAATAGCCTCACCCCTACATGTGGATAGAAGGAATGAAAAATAAT 600
| | | | |

Qy 1721 TGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAAGTACAAATCCATGAA 1780
| | | | |
Db 601 TGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAAGTACAAATCCATGAA 660
| | | | |

Qy 1781 AAGCTCAC 1788
| | | | |
Db 661 AAGCTCAC 668
| | | | |

RESULT 7
CF110345 1138 bp mRNA linear EST 23-JUL-2003
LOCUS Shultzomica03596 Rat lung airway and parenchyma cDNA libraries
DEFINITION Rattus norvegicus cDNA clone Contig3181 5', mRNA sequence.
ACCESSION CF110345
VERSION CF110345.1 GI:33166111
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1138)
Shultz,M.A., Zhang,L., Gu,Y.-Z., Baker,G.L., Fannuchi,M.V.,
Padua,A.M., Gurske,W.A., Morin,D., Penn,S.G., Jovanovich,S.B.,
Flopper,C.G. and Buckpitt,A.R.
Gene Expression Analysis in Response to Lung Toxicants: I.
Sequencing and Microarray Development
Unpublished (2003)
Contact: Shultz MA
Dept. of Molecular Biosciences, School of Veterinary Medicine
University of California, Davis
1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
Tel: 530 752 0793
Fax: 530 752 4698
Email: mashultz@ucdavis.edu
Average Phred score is 20 or better. All poor quality data (Phred <
20) and vector/linker sequence has been removed.
High quality sequence stop: 1138.
FEATURES
Location/Qualifiers
1..1138
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="Contig3181"
/sex="male"
/tissue_type="airway or parenchyma"
/dev_stage="adult"
/clone_lib="Rat lung airway and parenchyma cDNA libraries"
/note="Organ: lung; Vector: pGEM-11zf(-); Site 1: Eco RI;
Site 2: Not I; mRNA was isolated from microdissected rat
lung airways and parenchyma tissues."
ORIGIN
Query Match 30.5%; Score 647.2; DB 14; Length 1138;
Best Local Similarity 78.4%; Pred. No. 4.4e-124;
Matches 844; Conservative 0; Mismatches 213; Indels 19; Gaps 5;

Qy 137 GGGCTGGCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTA 196
| | | | |
Db 14 GGGTCTGGCGGCTGCATAGCCGCCACTGGAATGGACATGTGGAGCCTCAGGACCTGTA 73
| | | | |

Qy 197 CGACAAACCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGAGCTGCGTGAGGCA 256
| | | | |
Db 74 TGATAACCCAGTCACCTCCGTGTTCCAGTATGAAGGGCTCTGGAGAGTTGCGTGCAACA 133
| | | | |

Qy 257 GAGTTCAGGCTTCACCGAATGCAGGCCCTATTTCACCATCTCTGGGAC-TTCCAGCCATGC 315
| | | | |
Db 134 GAGCTCGGGTTCACTGAGTGCCGGCCATACTTCAACATCTCTGGGGCTTCCAGCCATGC 193
| | | | |

Qy 316 TGCAGGCAGTGCAGCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTGGCCTCCTGG 375
| | | | |
Db 194 TGCAAGCAGTGCAGCCCTGATGATCGTGGGCATTGTTCTGGGGGTTCATCGGCATCCTGG 253
| | | | |

Qy 376 TATCCATCTTTGCCCTGAAATGCATCCGATTGGCAGCATGGAGGACTCTGCCAAAGCCA 435
| | | | |
Db 254 TGTCCATCTTCGCCCTGAAAGTCATTCCGATTGGCAGCATGGATGACTCTGCCAAGGCCA 313
| | | | |

Qy 436 ACATGACACTGACCTCCGGGATCATGTTCAATTGTTCTCAGGTCTTTGTGCAATTGCTGGAG 495
| | | | |
Db 314 AGATGACTCTGACTTCTGGGATTATGTTTCATCATCTCTCAGGAGTCTCTGCAATCATTTGGTG 373
| | | | |

QY 496 TGCTGTGTTGGCCAACTGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACA 555
Db 374 TGCTGTGTTGGCCAACTGCTGGTAACCAACTTCTGGATGTCCACGGCCAACTGTACA 433
QY 556 CCGGCATGGTGGGATGGTGACAGACTGTTTCAGACCAGGTACACATTTGGTGGGCTCTGT 615
Db 434 GCGCATGGTGGGATGGTGACAGCCGTTTCAGACCAGGTACACTTTTGGTGGAGCTCTGT 493
QY 616 TCGTGGGCTGGTGGTGGGCTGCTGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCT 675
Db 494 TTGTGGGCTGGATCGCTGGAGCCCTCACACTAATTGGAGGTGTGATGATGTGCATCGCTT 553
QY 676 GCCGGGCTGGCCACAGAAAGAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCC 735
Db 554 GCCGTGGCCTGACTCCTGATGACCGCAACTTCAAAAGCCGTGTCTACACGCTTCTGGCC 613
QY 736 ACAGTGTGCTTACAAAGCCTGGAGGCTTCAAGCCAGCACTGGCTTTGGTCCCAACACCA 795
Db 614 AAAATGTTGCCATACAAAGCCTGGAGGCTTTAAGCCAGCACTGGCTTTGGTCCCAACGCCA 673
QY 796 AAAACAAGAGATATACGATGGAGGTGCCCGCACAGAGCAGGTACAACTTATCCTT 855
Db 674 GAAACAAGAGATATATGATGGGGTGCCTGACAGAGCAGGTGAACAATCTCATCCTA 733
QY 856 CCAAGCAGCACTATGTGTAATGCTCTAAGACCTCTCAGC-----ACGGCGGAAGAA- 908
Db 734 CCAAGTACGACTATGTGTAGTGTCTAAGACCGCCCCCAACATGTGTGAGAGAAAC 793
QY 909 -----CTCCCGGAGCTCACCCCAAAAACAAGGAGATCCCATCTAGATTCTTCTGCTT 964
Db 794 CCTTCCCCAAGGAAAGCTCACCCCCAAGCAAGGAGATCTACCTTGTTCCTTGTGCTT 853
QY 965 TTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTTTCATCTTTGGAGAGGCCAAATGGTC 1024
Db 854 TTGACTGACGGTTGAAAGTTGGTAAAGTCTTAATTTTCATCTGTGGGAGGTAGACAGTC 913
QY 1025 TTAGCCTCAGTCTCTGCTCTAAATATTTCCACCATAAAAACAGCTGAGTTATTTATGAATT 1084
Db 914 TTGGCCACA-----CGTCTCTAAATACTCCATCACAAAACAACCGAGTTACCAATTA--C 966
QY 1085 AGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTAAATATAAATTTCTACTC 1144
Db 967 AGGTTAAGAGACCACCAACCCACCCCTTCCCAATCCTCTGTTTTTTACCATAGTAAATC 1026
QY 1145 TGATGAGAGAATGTGGTTTAAATCTCTCTCTCACATTTTGATGATTTAGACAGACT 1200
Db 1027 TGGTGATGGAATGTAAATTCGACTCATGGACTAATACTTTGATAATTTCTAGAAATT 1082

RESULT 8
BM759274
LOCUS BM759274 680 bp mRNA linear EST 04-MAR-2002
DEFINITION K-EST0039301 S9SNU601 Homo sapiens CDNA clone S9SNU601-22-E05 5', mRNA sequence.
ACCESSION BM759274
VERSION BM759274.1 GI:19088889
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 680)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470

Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 22 row: E column: 05
High quality sequence stop: 680.
FEATURES
Location/Qualifiers
source
1..680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S9SNU601-22-E05"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/cell_line="SNU-601"
/lab_host="Top10F"
/clone_lib="S9SNU601"
/note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI; Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18S-FL3 vector. The obtained cDNA vectors were used for transporation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 29.4%; Score 623; DB 12; Length 680;
Best Local Similarity 100.0%; Pred. No. 5.1e-119;
Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1166 ATCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTCTTCTCCTCCTAGTCAATAA 1225
Db 1 ATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTCTTCTCCTCCTAGTCAATAA 60
QY 1226 CCATTTGATGATCTATTTCCAGCTTATCCCAAGAAAACCTTTTGAAGGAAAAGTAGA 1285
Db 61 CCATTTGATGATCTATTTCCAGCTTATCCCAAGAAAACCTTTTGAAGGAAAAGTAGA 120
QY 1286 CCAAAAGATGTTATTTTCTGCTGTTTGAATTTGTCTCCCCACCCCACTTGGCTAGTA 1345
Db 121 CCAAAAGATGTTATTTTCTGCTGTTTGAATTTGTCTCCCCACCCCACTTGGCTAGTA 180
QY 1346 ATAAACACTTACTGAAGAAGCAATAGAGAAAAGATATTTGTAATCTCTCCAGCCCAT 1405
Db 181 ATAAACACTTACTGAAGAAGCAATAGAGAAAAGATATTTGTAATCTCTCCAGCCCAT 240
QY 1406 GATCTCGGTTTTCTTACACTGTGATCTTAAAGTTACCAACCAAGTCATTTTCAGTTT 1465
Db 241 GATCTCGGTTTTCTTACACTGTGATCTTAAAGTTACCAACCAAGTCATTTTCAGTTT 300
QY 1466 GAGGCAACCAACCTTTCTACTGCTGTTGACATCTTATTACAGCAACACCATTTCTAG 1525
Db 301 GAGGCAACCAACCTTTCTACTGCTGTTGACATCTTATTACAGCAACACCATTTCTAG 360
QY 1526 GAGTTTCTGAGCTCTCCACTGGAGTCCCTTTTGTGCGGGTCAGAAATTTGCCCTAG 1585
Db 361 GAGTTTCTGAGCTCTCCACTGGAGTCCCTTTTGTGCGGGTCAGAAATTTGCCCTAG 420
QY 1586 ATGAATGAGAAAATTTATTTTAAATTTAACTCCTAAATATAGTTAAATAAATAATGT 1645
Db 421 ATGAATGAGAAAATTTATTTTAAATTTAACTCCTAAATATAGTTAAATAAATAATGT 480
QY 1646 TTATGATAAATGATACACTATCTCTGTGAATAGCCTCACCCCTACATGTGGATAGAAG 1705
Db 481 TTATGATAAATGATACACTATCTCTGTGAATAGCCTCACCCCTACATGTGGATAGAAG 540

QY 1706 AAATGAAAAAATAATTGCTTTGACATTTGCTATATGTTACTTTGTAAAGTCATGCTTAAG 1765
|||||
Db 541 AAATGAAAAAATAATTGCTTTGACATTTGCTATATGTTACTTTGTAAAGTCATGCTTAAG 600
|||||
QY 1766 TACAAATTCATGAAAGCTCAC 1788
|||||
Db 601 TACAAATTCATGAAAGCTCAC 623
|||||
RESULT 9
AY402518 795 bp DNA linear GSS 15-DEC-2003
LOCUS Mus musculus CLDN18 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402518
VERSION AY402518.1 GI:39758504
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 795)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 795)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..795
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>795
/gene="CLDN18"
/locus_tag="HCM1250"
ORIGIN
Query Match 28.6%; Score 607.6; DB 29; Length 795;
Best Local Similarity 86.4%; Pred. No. 8e-116;
Matches 686; Conservative 0; Mismatches 99; Indels 9; Gaps 1;
QY 90 ATGTCCACCACCATGCCAAGTGGTGGCGTTCTCTCTGTCATCCTGGGCTGGCCGGC 149
|||||
Db 1 ATGGCCACCACCACGTCGAGGTGGTAGGGCTTCTCTGTCCTCTGCTGGCTGGCCGGC 60
|||||
QY 150 TGATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAACCCCGTC 209
|||||
Db 61 TGATAGCCGCCACTGGGATGGACATGTGGAGCACTCAAGACCTGTATGACAACCCAGTC 120
|||||
QY 210 ACCTCGGTCCAGTACGAAGGGCTCTGGAGGAGCTGGTGGGCAGAGTTCAGGCTTC 269
|||||
Db 121 ACCGCGGTTCAGTATGAAGGGCTCTGGAGGAGTTGGTGCAACAGAGCTCGGGGTTTC 180
|||||
QY 270 ACCGAATCAGGCCCTATTTCAACCATCCTGGGACTTCAGCCATGCTGCAGGAGTGCGA 329
|||||
Db 181 ACCGAGTCCGGCCATACCTTCAACCATCCTGGGCTTCAGCCATGCTGCAAGCTGTACGA 240
|||||
QY 330 GCCCTGATGATCGTAGGCATCGTCTGGTGCCATTGGGCTCTGTTATCCATCTTTGCC 389
|||||
Db 241 GCCCTGATGATCGTGGGCATTGTTCTGGGGGTGTCATCGGTATCCTCGTGTCCATCTTCGCC 300
|||||

QY 390 CTGAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACACTGACC 449
|||||
Db 301 CTGAAGTGCATTCGCATTGGTAGCATGGATGACTCTGCCAAGGCCAAGATGACTCTGACT 360
|||||
QY 450 TCCGGGATCATGTTCTCAGGTCTTTGTGCAATTCCTGGAGTGTCTGTGTTTGGC 509
|||||
Db 361 TCTGGGATCTTGTTCATCATCTCCGGCATCTGTGCAATCATTTGGTGTCTGTGTTTGGC 420
|||||
QY 510 AACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAAACATGTACA-----CCGGC 560
|||||
Db 421 AACATGCTGGTGACCAACTTCTGGATGTCCACAGCTAAACATGTACAGCGGCATGGCGGC 480
|||||
QY 561 ATGGGTGGGATGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGGCTCTGTTCGTG 620
|||||
Db 481 ATGGGTGGCATGGTGCAGACCCGTTTCAGACCAGGTACACCTTTGGTGCAGCTCTGTTCGTG 540
|||||
QY 621 GGCTGGTCTGCTGGAGCCCTCACACTAAATTGGGGGTGTGATGATGTGCATCGCCTGCCGG 680
|||||
Db 541 GGCTGGGTGCTGGAGCCCTCACCTGATTTGGGGAGTGTGATGTGCATCGCCTGCCGT 600
|||||
QY 681 GGCTGGCACCAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCCTCAGGCCACAGT 740
|||||
Db 601 GGCTGACACCATGACAGCAACTTCAAAGCTGTGTCTTACCATGCTCTGGCCAAAT 660
|||||
QY 741 GTTGCTTACAAGCCTGGAGGCTTCAAGCCAGCACTGGCTTTGGTCCAACACCAAAAC 800
|||||
Db 661 GTTGCTTACAGGCTGGAGGCTTTAAGGCCAGCACTGGCTTTGGTCCAACACCAAGAAC 720
|||||
QY 801 AAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCTTCCAAG 860
|||||
Db 721 AAGAAGATCTACGATGGGGGTGCCCGCACAGAGACGATGAACAGTCTCATCTACCAAG 780
|||||
QY 861 CACGACTATGTGTA 874
|||||
Db 781 TATGACTATGTGTA 794
|||||
RESULT 10
AY402517 617 bp DNA linear GSS 15-DEC-2003
LOCUS Pan troglodytes CLDN18 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402517
VERSION AY402517.1 GI:39758503
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
1 (bases 1 to 617)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 617)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..617
/organism="Pan troglodytes"
/mol_type="genomic DNA"

```

gene
/db_xref="taxon:9598"
<!.>617
/gene="CLDN18"
/locus_tag="HCM1250"

ORIGIN
Query Match      28.0%; Score 593.2; DB 29; Length 617;
Best Local Similarity 96.4%; Pred. No. 8.2e-113;
Matches 595; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 259 GTTCAGGCTTCACCGAATGAGGCCCTATTTCCACATCCTGGGACTTCCAGCATGCTGC 318
    |||||
Db 1 GTTCAGGCTTCACCGAATGAGGCCCTATTTCCACATCCTGGGCTTCNNCCATGCTGC 60

QY 319 AGGCAGTGCAGCCCTGATGATCGTAGGCATCGTCCCTGGGTGCCATTTGGCTCCTGGTAT 378
    |||||
Db 61 AGGCAGTGCAGCCCTGATGATCGTAGGCATCGTCCCTGGGTGCCATTTGGCTCCTGGTAT 120

QY 379 CCATCTTTGCCCTGAAATGCATCCGCATTTGGCAGCATGGAGACTCTGCCAAAGCCAACA 438
    |||||
Db 121 CCATCTTTGCCCTGAAATGCATCCGCATTTGGCAGCATGGAGACTCTGCCAAAGCCAACA 180

QY 439 TGACACTGACCTCCGGGATCATGTTTCAATGTTCTCAGGTCTTTGTGCAATTCCTGGAGTG 498
    |||||
Db 181 TGACACTGACCTCCGGGATCATGTTTCAATGTTCTCAGGTCTTTGTGCAATTCCTGGAGTG 240

QY 499 CTGTGTTTGCCAAACATGCTGGTGACTAACTTCTGGATGTCACAGCTAACATGTACACCG 558
    |||||
Db 241 CTGTGTTTGCCAAACATGCTGGTGACTAACTTCTGGATGTCACAGCTAACATGTACACCG 300

QY 559 GCATGGGTGGGATGGTGACAGACTGTTTCAGACCAAGGTACACATTTGGTGGGCTCTGTTGC 618
    |||||
Db 301 GCATGGGTGGGATGGTGACAGACTGTTTCAGACCAAGGTACACATTTGGTGGGCTCTGTTGC 360

QY 619 TGGGCTGGTGCCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTCATCGCCTGCC 678
    |||||
Db 361 TGGGCTGGTGCCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTCATCGCCTGCC 420

QY 679 GGGGCTGGCACCAAGAAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACA 738
    |||||
Db 421 GGGGCTGGCACCAANNNNNNNNNNCTACAAAGCCGTTTCTTATCATGCTCGGGCCACA 480

QY 739 GTGTTGCCTACAAGCTGGAGGCTTCAAGGCCAGCAGCTGGCTTTGGTCCACACCAAAA 798
    |||||
Db 481 GTGTTGCCTACAAGCTGGAGGCTTCAAGGCCAGCAGCTGGCTTTGGATCCACACCAAAA 540

QY 799 ACAAGAAGATATACGATGGAGGTGCCCGGCACAGAGGACGAGGTACATCTTATCCTTCCA 858
    |||||
Db 541 ACAAGAAGATATNNATGGAGGTGCCCGNACNNAGGACGAGGTACAATCTTATCCTTCCA 600

QY 859 AGCAGCACTATGTGTAA 875
    |||||
Db 601 AGCAGCACTATGTGTAA 617

RESULT 11
BQ016897
LOCUS
DEFINITION
  UI-H-D10-auz-b-03-0-UI.s1 NCI_CGAP_D10 Homo sapiens cDNA clone
  IMAGE:5876258 3', mRNA sequence.
ACCESSION
  BQ016897
VERSION
  BQ016897.1 GI:19752174
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 627)
AUTHORS
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished (1997)
COMMENT
  Contact: Robert Strausberg, Ph.D.
```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
sequence: 1-39, >AT-rich#Low complexity (matched complement)
500-562, >AT-rich#Low complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

1..627
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5876258"
/tissue_type="Lung Focal Fibrosis"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_D10"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP D10 is a cDNA library containing the following
tissue(s): A pool of Lung Focal Fibrosis. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
ATACGCGGTC.
TAG_TISSUE=lung with fibrosis
TAG_LIB=UI-H-D10
TAG_SEQ=ATACGCGGTC"

ORIGIN

Query Match 27.7%; Score 588.2; DB 12; Length 627;
Best Local Similarity 97.7%; Pred. No. 8.9e-112;
Matches 607; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 1102 TTTTCAATCCTCTATTTCTTTTAAATATAAATTTCTACTCTGATGAGAGATGTGGT 1161
 |||||
Db 7 TTTTATTTTAAATTTATTTTAAATATAAATTTCTACTCTGATGAGAGATGGGT 66

QY 1162 TTTAATCTCTCTCACATTTTGTATGATTTAGACAGACTCCCCCTCTTCTCCTAGTCAA 1221
 |||||
Db 67 TTTAATCTCTCTCACATTTTGTATGATTTAGACAGACTCCCCCTCTTCTCCTAGTCAA 126

QY 1222 TAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAACTTTTGAAAGGAAGAG 1281
 |||||
Db 127 TAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAACTTTTGAAAGGAAGAG 186

QY 1282 TAGACCCAAAGATGTTATTTCTGCTGTTGAATTTTGTCTCCACCCCCCAACTTGGCT 1341
 |||||
Db 187 TAGACCCAAAGATGTTATTTCTGCTGTTGAATTTTGTCTCCACCCCCCAACTTGGGT 246

QY 1342 AGTAATAAACACTTACTGAAGAAGAACCAATAAGAGAAAGATATTGTAATCTCTCCAGC 1401
 |||||
Db 247 AGTAATAAACACTTACTGAAGAAGAACCAATAAGAGAAAGATATTGTAATCTCTCCAGC 306

QY 1402 CCATGATCTCGGTTTTCTTACACTGTGATCTTAAAGTTACCAACCAAGTCATTTTCA 1461
 |||||
Db 307 CCATGATCTCGGTTTTCTTACACTGGGATCTTAAAGTTACCAACCAAGTCATTTTCA 366

QY 1462 GTTTGAGGCAACCAACCTTTCTACTGCTGTTGACATCTTCTTATTACAGCAACCACTT 1521
 |||||
Db 367 GTTTGAGGCAACCAACCTTTCTACTGCTGTTGACATCTTCTTATTACAGCAACCACTT 426

QY 1522 CTAGGAGTTTCTCGAGCTCTCCACTGGAGTCCTCTTTCTGCGGGTCCAGAAATTGTCC 1581
|||||
Db 427 CTAGGAGTTTCTCGAGCTCTCCACTGGAGTCCTCTTTCTGCGGGTCCAGAAATTGTCC 486
|||||
QY 1582 CTAGATGAATGAGAAATTA-TTTTITTTTAAATTAACTCCTAAATATAGTTAAATAAAT 1640
|||||
Db 487 CTAGATGAATGAGAAATTA-TTTTITTTTAAATTAACTCCTAAATATAGTTAAATAAAT 546
|||||
QY 1641 AATGTTTTAGTAAATGATACACTATCTGTGAAATAGCCTCACCCCTACATGTGGATA 1700
|||||
Db 547 AATGTTTTAGTAAATGATACACTATCTGTGAAATAGCCTCACCCCTACATGTGGATA 606
|||||
QY 1701 GAAGGAAATGAAATAAATT 1721
|||||
Db 607 GAAGGAAATGAAATAAATT 627
|||||

RESULT 12
BI823718
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BI823718 743 bp mRNA linear EST 04-OCT-2001
603040838F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5181825 5',
mRNA sequence.
BI823718 GI:15935268
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 743)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11453 Row: k column: 10
High quality sequence stop: 554.

FEATURES
source
1. .743
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5181825"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 26.0%; Score 552; DB 12; Length 743;
Best Local Similarity 95.4%; Pred. No. 3e-104;
Matches 666; Conservative 0; Mismatches 20; Indels 12; Gaps 9;
QY 28 CACACCTTCGGCAGCAGGAGGGCGGAGCTTCTCGAGCGGCGCAGGCGGGCCAGGA 87
|||||
Db 1 CACACCTTCGGCAGCAGGAGGGCGGAGCTACTCGAGCGGCGCAGGCGGGCCAGGA 60
|||||
QY 88 TCATGTCCACCACCACATGCCAAAGTGGTGGCGTTCTCTGTCCATCTCTGGGCTGGCCG 147
|||||

Db 61 TCATGTACAACACCACATGCCAAGTGGTGGCTTCTCTGTCCATCTGGGCTGGCCG 120
|||||
QY 148 GCTGCATCGCGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACCAACCCCG 207
|||||
Db 121 GCTGCATCGCGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACCAACCCCG 180
|||||
QY 208 TCACCTCCGTGTTCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCT 267
|||||
Db 181 TCACCTCCGTGTTCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCT 240
|||||
QY 268 TCACCGAATGCAGGCCCTATTTCACCATCTCTGGGACCTTCCAGCCATGCTGCAGGCAGTGC 327
|||||
Db 241 TCACCGAATGCAGGCCCTATTTCACCATCTCTGGGACCTTCCAGCCATGCTGCAGGCAGTGC 300
|||||
QY 328 GAGCCCTGATGATCGTAGGCATCGTCTGGGTGCCA-TTGGCCTCTCTGGTATCCATCTTT 386
|||||
Db 301 GAGCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTGGCCTCTCTGGTATCCATCTTT 360
|||||
QY 387 GCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGAC-ACT 445
|||||
Db 361 GCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACAACT 420
|||||
QY 446 GACCTCCGGGATCATGTTTCATTTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTT 505
|||||
Db 421 GACCTCCGGGATCATGTTTCATTTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTA 480
|||||
QY 506 TGCCAAACATGCTGTGACTAACTTCTGGATGTCCACAGCTAACTATACACCGGCATGGG 565
|||||
Db 481 TGCCAAACATGCTGTGACTAACTTCTGGATGTCCACAGCTAACTATGACACCGGCATGGG 540
|||||
QY 566 T-GGGATGGTGCAG-ACTGTTTCAGACCAGGTACACACTTTGGTGGCGC--TCTGTTCTGTGG 621
|||||
Db 541 TGGGATGGTGCAGAACTGTACAGACCAGGTACACACTTATAGGTGGCCATCTGTACGTGG 600
|||||
QY 622 GCTGGTGGCTGG-AGGCCTCACACTAATTGGGGGTGTGATG-ATGTGCATCG--CCTGC 677
|||||
Db 601 GCTGGTGGCTGGAGGACTCACACTAATTGAGGGTGTGATGATGATGTCATGTCATCGCCTGCC 660
|||||
QY 678 CGGGGCTGGCACCCAGAA--GAAACCAACTACAAAGCC 713
|||||
Db 661 GGGGGCTGTACACCAGAAAGAAACCAACTACAAAGCC 698
|||||

RESULT 13
CB462343
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES

CB462343 681 bp mRNA linear EST 26-MAR-2003
722418 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
CB462343
CB462343.1 GI:29268727
EST.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
1 (bases 1 to 681)
Smith,T.P.L., Roberts,A.J., Echterkamp,S.E., Chitko-McKown,C.G.,
Wray,J.E. and Keele,J.W.
A second set of bovine ESTs from pooled-tissue normalized libraries
Unpublished (2003)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: LAM8001 row: I column: 12
Seq primer: GTAATACGACTCATATAGG.
Location/Qualifiers

```
source 1..681
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

ORIGIN
Query Match 24.8%; Score 525.8; DB 14; Length 681;
Best Local Similarity 86.9%; Pred. No. 8.5e-99;
Matches 578; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 44 GGAGGGCGGCAGCTTCTCGAGCGCGGCAGGGCGGGCCAGGATCATGTCCACCACCAC 103
Db 17 GGCAGGACGAAGCGCGAGCTGGCGGGAGGACGGAGAGCCGGCCATGTCCACCACCAG 76

Qy 104 ATGCCAAGTGGTGGCTTCTCTGTCCATCCTGGGGCTGGCGGCTGCATCGCGGCCAC 163
Db 77 ATGCCAAGTGGTGGCTTCTCTGTCCATCCTGGGGCTGGCGGCTGCATCGTGGCCAC 136

Qy 164 CGGATGGACATGTGGAGCACCCAGGACCTGTACGACAAACCCCGTCACCTCCGTGTCCA 223
Db 137 GGAGATGGACATGTGGAGCACCCAGGACCTATACGACAAACCCGTCACCGCTGTGTCCA 196

Qy 224 GTACGAAGGCTCTGGAGGAGCTGCGTGAGGCGAGAGTTCAGGCTTCACCGAATGCAGGCC 283
Db 197 GTATGAAGGCTCTGGCGCAGTTGCGTGCAGCAGAGCTCAGGCTTCACCGAGTCCGGCC 256

Qy 284 CTATTTACCATCTCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATCGT 343
Db 257 CTACCTCACCATCTCTGGGCTGCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATCGT 316

Qy 344 AGGCATCGTCTGGGTGCCATTTGGCCTCCTGTATCCATCTTTGGCCCTGAATGCATCCG 403
Db 317 GGGCATCGTCTGAGTGTCAATGGCTCCTGGTGGCCATCTTTGGCCCTGAATGCATCCG 376

Qy 404 CATTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACACTGACCTCGGGATCATGTT 463
Db 377 TATGGCAACATGGATGACTCGGCCAAAGCCAAATGACACTCACCCTCGGGATCATGTT 436

Qy 464 CATGTCTCAGGTCTTTGTGCAATTGCTGGAATGCTGTGTTTGCACATGCTGTTGAC 523
Db 437 CATCATCGCAGGTCTCTGTGCAATCGTGGAGTGTGTGTTTGCACATGCTGTTTAC 496

Qy 524 TAATCTCTGGATGTCACAGCTAACATGTACACCGGCATGGTGGGATGGTGCAGACTGT 583
Db 497 TAATCTCTGGATGTCACAGCCAGCATGTTTACAGCATGGGGGGATGGTGCAGACCGT 556

Qy 584 TCAGACCAGGTACATATTGGTGGGCTCTGTTCGTGGGCTGGTTCGTGGAGGCCTCAC 643
Db 557 TCAGACCAGGTACATTTGGTGGGCTCTGTTCGTGGGCTGGTTCGTGGAGGCCTCAC 616

Qy 644 ACTAATTGGGGTGTGATGTGATCGCATCGCCTGCCGGGCTGGCCACGAGAAACCAA 703
Db 617 GCTGATTGGGGCGTGTGATGTGATCGCATCGCCTGCCGGGCTGTGGCCCGCCGAGAAACCAA 676

Qy 704 CTACA 708
Db 677 CTACA 681

RESULT 14
BM837765
LOCUS
DEFINITION K-EST0113903 S9SNU601 Homo sapiens cDNA clone S9SNU601-77-B11 5',
mRNA sequence.
ACCESSION BM837765
VERSION BM837765.1 GI:19194174
```

```
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 651)
Oh,K.J., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Kim,N.S., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 77 row: B column: 11
High quality sequence stop: 651.
Location/Qualifiers
1..651
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S9SNU601-77-B11"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/lab_host="SNU-601"
/clone_lib="S9SNU601"
/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested pME18S-FL3 vector. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
```

ORIGIN

```
Query Match 24.5%; Score 519; DB 12; Length 651;
Best Local Similarity 100.0%; Pred. No. 2.2e-97;
Matches 519; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1270 GAAAGGAAAGAGTAGACCCAAAGATGTTATTTCTGCTGTTTGAATTTGTCTCCCACC 1329
Db 1 GAAAGGAAAGAGTAGACCCAAAGATGTTATTTCTGCTGTTTGAATTTGTCTCCCACC 60

Qy 1330 CCCAACTTGGCTAGTAATAAAACACTTACTGAAGAAGAAAGCAATAAGAGAAAGATATTGT 1389
Db 61 CCCAACTTGGCTAGTAATAAAACACTTACTGAAGAAGAAAGCAATAAGAGAAAGATATTGT 120

Qy 1390 AATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAGTTACCAACCA 1449
Db 121 AATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAGTTACCAACCA 180

Qy 1450 AAGTCATTTTCAGTTGAGGCAACCAACCTTTCTACTGCTGTGACATCTTCTTATTAC 1509
Db 181 AAGTCATTTTCAGTTGAGGCAACCAACCTTTCTACTGCTGTGACATCTTCTTATTAC 240

Qy 1510 AGCAACACCATTTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTCTTTCTGTGCGGGT 1569
Db 241 AGCAACACCATTTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTCTTTCTGTGCGGGT 300
```


FEATURES
source
High quality sequence stop: 549.
Location/Qualifiers
1. .549
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNU520-20-C12"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNU520"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 20.5%; Score 434; DB 12; Length 549;
Best Local Similarity 100.0%; Pred. No. 1.1e-79;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1355 TACTGAAGAAGCAATPAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGT 1414
Db 1 TACTGAAGAAGCAATPAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGT 60
Qy 1415 TTCTTTACACTGTGATCTTAAAGTTACCAACCAAGTCATTTTCAGTTTGAGGCAACC 1474
Db 61 TTCTTTACACTGTGATCTTAAAGTTACCAACCAAGTCATTTTCAGTTTGAGGCAACC 120
Qy 1475 AAACCTTTTCTACTGCTGTGACATCTTCTTATTACAGCAACACCATTTCTAGGAGTTTCCT 1534
Db 121 AAACCTTTTCTACTGCTGTGACATCTTCTTATTACAGCAACACCATTTCTAGGAGTTTCCT 180
Qy 1535 GAGCTCTCCACTGGAGTCCTCTTTCTGTGCGGGTCAGAAATTGTCCTTAGATGAATGAG 1594
Db 181 GAGCTCTCCACTGGAGTCCTCTTTCTGTGCGGGTCAGAAATTGTCCTTAGATGAATGAG 240
Qy 1595 AAAATTATTTTTTTTAAATTAAGTCCTAAATATAGTTAAATAAATAATATGTTTTAGTAA 1654
Db 241 AAAATTATTTTTTTTAAATTAAGTCCTAAATATAGTTAAATAAATAATATGTTTTAGTAA 300
Qy 1655 ATGATACACTATCTCTGTGAATAGCCTCACCCCTACATGTGGATAGAGGAAATGAAAA 1714
Db 301 ATGATACACTATCTCTGTGAATAGCCTCACCCCTACATGTGGATAGAGGAAATGAAAA 360
Qy 1715 AATAATTGCTTTGACATTGTCTATATGGTACTTTTGTAAGTCAATGCTTAAGTACAAATTC 1774
Db 361 AATAATTGCTTTGACATTGTCTATATGGTACTTTTGTAAGTCAATGCTTAAGTACAAATTC 420
Qy 1775 CATGAAAAGCTCAC 1788
Db 421 CATGAAAAGCTCAC 434

RESULT 17
BG538299 961 bp mRNA linear EST 03-APR-2001
LOCUS 602566769P1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4691417 5',
DEFINITION mRNA sequence.

ACCESSION BG538299
VERSION BG538299.1 GI:13530531
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 961)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1510 row: i column: 18
High quality sequence stop: 578.

FEATURES
source
Location/Qualifiers
1. .961
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4691417"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_77"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccgctcgcc); Site 2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGGGCGGCACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 19.5%; Score 413.2; DB 12; Length 961;
Best Local Similarity 86.0%; Pred. No. 2.2e-75;
Matches 626; Conservative 0; Mismatches 73; Indels 29; Gaps 14;
Qy 1028 GCCTCAGTCTCTGTCTCTAAATATTCACCAFAAAACAGCTGAGTTATTTATGAATTAGA 1087
Db 1 GCCTCAGTCTCTGTCTCTAAATATTCACCAFAAAACAGCTGAGTTATTTATGAATTAGA 60
Qy 1088 GGCATATAGCTCACATTTTCAATCTCTATTTCTTTTAAATATAACTTTTACTCTGA 1147
Db 61 GGCATATAGCTCACATTTTCAATCTCTATTTCTTTTAAATATAACTTTTACTCTGA 120
Qy 1148 TGAGAGAAATGTGGTTTT-----AATCTCTCTCTCACATTTTGTGATGATTAGACAGACTCC 1202
Db 121 TGAGAGTAATGTGGCTTTCTACACTCTCTCTCTCACATTTTGTGATGATTAGACAGACTCC 180
Qy 1203 CCCTCTTCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAA 1262
Db 181 CCCTCTTCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAA 240
Qy 1263 AACTTTTGTAAA-GGAAAGAGTAGAGCCCAAGATGTT-ATTTTCTGCTGTT-TGAATTTTG 1319
Db 241 AACTTTTGTAAACGGAAGAGTAGAGCCCAAGATGTTTCATTTTCTGCTGTTTCTGAATTTGT 300
Qy 1320 TCTCCCCACCCCAACTGGCTAGTAATA-AACACTTACTGAAGAGAAAGCAAT-AAGAG 1377
Db 301 ACTCCCCAGCCCCCACTTGGCTAGTAATAACAACACTTACTGAAGAGAAAGCAATCAGCAG 360
Qy 1378 AAAGATATTTGTAAATCTCTCCAGCCCATGATCTCGGTTTTTCTTACACTGTGATCTT-AAA 1436
Db 361 AAAGATATTTGTAAATCTCTCCAGCCCATGATCTCGGTTTTTCTTACACTGTGATCTTCAA 420

No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 17.6%; Score 373.6; DB 10; Length 421;
Best Local Similarity 97.0%; Pred. No. 4.2e-67;
Matches 391; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
QY 924 CCCAAAAACAGGAGATCCCATCTAGATTCTTCTGTTTACTCACAGCTGGAAGT 983
Db 12 CACCCAAAAACAGAGATCCCATCTAGATTCTTCT--GCTTTGACTCACAGCTGGAAGT 69
QY 984 TAGAAAAGCCTCGATTTCATCTTTGGAGAGGCCAAATGGTCTTAGCTCAGTCTCTGCT 1043
Db 70 TAGAAAAGCCTCGATTTCATCTTTGGAGAGGCCAAATGGTCTTAGCTCAGTCTCTGCT 129
QY 1044 CTAAATATTCACCAATAAAACAGCTGAGTTTATTTATGAATTAGAGGCTATAGCTCACATT 1103
Db 130 CTAAATATTCACCAATAAAACAGCTGAGTTTATTTATGAATTAGAGGCTATAGCTCACATT 189
QY 1104 TTCAATCCTCTATTCTTTTAAATATAAATTCTACTCTGATGAGAGAATGTGTTT 1163
Db 190 TTCAATCCTCTATTCTTTTAAATATAAATTCTACTCTGATGAGAGAATGTGTTT 249
QY 1164 TAATCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTCTTCTCCTAGTCAATA 1223
Db 250 TAATCTCTCTCACATTTNGATGATTTAGACAGACTCCCCCTCTTCTCCTAGTCAATA 309
QY 1224 AACCCATTGATGATCTAATTTCCAGCTTATCCCAAGAAAACCTTTGAAAGGAAAGAGTA 1283
Db 310 AACCCATTGATGATCTAATTTCCAGCTTATCCCAAGAAAACCTTTGAAAGGAAAGAGTA 369
QY 1284 GACCCAAAGATGTTTATTTCTGCTGTTTGAATTTTGTCTCCCC 1326
Db 370 GACCCAAAGATGTTTATTTCTGCTGTTTGGATTGTTGTTCTCCCC 412

RESULT 24
BE841291/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BE841291
MR0-SN0040-310500-007-all SN0040 Homo sapiens cDNA, mRNA sequence.
BE841291
BE841291.1 GI:10273669
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 379)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR0-SN0040-310

TITLE

JOURNAL
MEDLINE
PUBMED
COMMENT

FEATURES
source

500-007-all&t3=2000-05-31&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 379.
Location/Qualifiers
1..379
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="SN0040"
/note="Organ: stomach normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 17.0%; Score 360.8; DB 10; Length 379;
Best Local Similarity 99.5%; Pred. No. 1.9e-64;
Matches 362; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1211 CTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCAAGAAAACCTTTG 1270
Db 367 CGCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCAAGAAAACCTTTG 308
QY 1271 AAAGGAAAGAGTAGACCCAAAGATGTTATTTCTGCTGTTGAATTTGTCTCCCCACCC 1330
Db 307 AAAGGAAAGAGTAGACCCAAAGATGTTATTTCTGCTGTTGAATTTGTCTCCCCACCC 248
QY 1331 CCAACTTGGCTAGTAATAAACACTTACTGAGAAGAAAGCAATAAGAGAAGATATTGTA 1390
Db 247 CCAACTTGGCTAGTAATAAACACTTACTGAGAAGAAAGCAATAAGAGAAGATATTGTA 188
QY 1391 ATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATCTTAAAGTTACCAACCAA 1450
Db 187 ATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATCTTAAAGTTACCAACCAA 128
QY 1451 AGTCATTTTCAGTTTGAGGCAACCAACCTTCTACTGCTGTTGACATCTTCTATTACA 1510
Db 127 AGTCATTTTCAGTTTGAGGCAACCAACCTTCTACTGCTGTTGACATCTTCTATTACA 68
QY 1511 GCAACACCATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCCTTCTTCTGCGGGTC 1570
Db 67 GCAACACCATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCCTTCTTCTGCGGGTC 8
QY 1571 AGAA 1574
Db 7 AGAA 4

RESULT 25

N99013

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

N99013
za67a07.r1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
IMAGE:297588 5', mRNA sequence.
N99013
N99013.1 GI:1270438
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 420)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfsing,T., Soares,M., Tan,F.,
Trevisan,E., Waterston,R., Williamson,A., Wohldmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK

Db 302 TGCTGTTGACATCTTCTTATTACAGCAACACCATTTCTAGAGTTTCTCTGAGCTCTCCACT 361

QY 1547 GGAGTCCTCTTTCTGTCGGGGTCCAGAAATTGTCCCTAGATGAATGAGAAAATTA-TTTT 1605
|||||

Db 362 GGAGTCCTCTTTCTGGCGGGTTCAGAAAATGGCCCTAGATGAATGAGAAAATTTATTTT 421
|||||

QY 1606 TTTTAATTTAAGTCCTAAATATAGTTAAATAAA 1639
|||||

Db 422 TTTTAATTTAAGGCCCAATATAGTTTCAATAAA 455
|||||

RESULT 27

BM836109

LOCUS

DEFINITION K-EST0111603 S9SNU601 Homo sapiens cDNA clone S9SNU601-65-G03 5',
mRNA sequence.

ACCESSION BM836109

VERSION BM836109.1 GI:19192518

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.

REFERENCE

AUTHORS

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE

21C Frontier Korean EST Project 2001

JOURNAL

Unpublished (2002)

COMMENT

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 65 row: G column: 03

High quality sequence stop: 341.

Location/Qualifiers

1. .341

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S9SNU601-65-G03"

/sex="M"

/tissue_type="Ascites"

/cell_type="Epithelial"

/cell_line="SNU-601"

/lab_host="Top10F"

/clone_lib="S9SNU601"

/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested pME18S-FL3 vector. The
obtained cDNA vectors were used for transporation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 16.1%; Score 341; DB 12; Length 341;

Best Local Similarity 100.0%; Pred. No. 2.6e-60;

Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 874 AATGCTCTAAGACCTCTCAGCAGCGCGGAAGAACTCCGGAGAGCTACCCAAAAAC 933
|||||

Db 1 AATGCTCTAAGACCTCTCAGCACGGCGGAAGAAACTCCCGAGAGCTCACCCAAAAAC 60

QY 934 AAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTAGAAAAAGCC 993
|||||

Db 61 AAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTAGAAAAAGCC 120
|||||

QY 994 TCGATTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTCTCTCTAAATATTC 1053
|||||

Db 121 TCGATTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTCTCTAAATATTC 180
|||||

QY 1054 CACCATAAAACAGCTGAGTTATTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTC 1113
|||||

Db 181 CACCATAAAACAGCTGAGTTATTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTC 240
|||||

QY 1114 TATTTCTTTTTTAAATATAACTTTTCTACTCTGATGAGAGAAATGGTTTAACTCTCT 1173
|||||

Db 241 TATTTCTTTTTTAAATATAACTTTTCTACTCTGATGAGAGAAATGGTTTAACTCTCT 300
|||||

QY 1174 CTCACATTTTGATGATTTAGACAGACTCCCTCTTCTCTCC 1214
|||||

Db 301 CTCACATTTTGATGATTTAGACAGACTCCCTCTTCTCTCC 341
|||||

RESULT 28

AW867300/c

LOCUS

DEFINITION MR3-SN0036-280400-009-f08 SN0036 Homo sapiens cDNA, mRNA sequence.

ACCESSION AW867300

VERSION AW867300.1 GI:8001352

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 412)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,P., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

PUBMED

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR3-SN0036-280
400-009-f08&t3=2000-04-28&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 5

High quality sequence stop: 412.

Location/Qualifiers

1. .412

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="SN0036"

/note="Organ: stomach normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of

FEATURES

source

| | | | |
|-------------------------|---|---|-----------------------|
| Seq primer: M13 FORWARD | | AUTHORS | |
| POLYA=Yes. | | Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S. | |
| FEATURES | Location/Qualifiers | | TITLE |
| | 1. .493 | | |
| SOURCE | /organism="Homo sapiens" | | JOURNAL |
| | /mol_type="mRNA" | | |
| FEATURES | /db_xref="taxon:9606" | | COMMENT |
| | /clone="UI-H-DT0-aud-1-05-0-UI" | | |
| FEATURES | /tissue_type="Metastatic Chondrosarcoma" | | FEATURES |
| | /dev_stage="Adult" | | |
| FEATURES | /lab_host="DH10B (Life Technologies)" | | source |
| | /clone_lib="NCI_CGAP DT0" | | |
| FEATURES | /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI CGAP DT0 is a cDNA library containing the following tissue(s): Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag (dT)18 tail. The sequence tag for this library is AACTGTTCCG. | | ORIGIN |
| | TAG_TISSUE=lung metastatic chondrosarcoma | | |
| ORIGIN | TAG_LIB=UI-H-DT0 | | Query Match |
| | TAG_SEQ=AACTGTTCCG" | | |
| ORIGIN | 16.0%; Score 339; DB 14; Length 493; | | Best Local Similarity |
| | 100.0%; Pred. No. 6.5e-60; | | |
| ORIGIN | Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | Matches |
| | | | |
| ORIGIN | 1783 GCTCACACCTGTATCCTAGCACCTTTGGAGGCTGAGGAGGAGGATCATTGAGCCCCAG 1842 | | QY |
| | 355 GCTCACACCTGTATCCTAGCACCTTTGGAGGCTGAGGAGGAGGATCATTGAGCCCCAG 296 | | |
| ORIGIN | 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAA 1902 | | QY |
| | 295 AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAA 236 | | |
| ORIGIN | 1903 AAAATCAGCCAGTCAATGCTGGCATAACACCTGTAGTCCCAGCATTCGCGGAGGCTGAGGTG 1962 | | QY |
| | 235 AAAATCAGCCAGTCAATGCTGGCATAACACCTGTAGTCCCAGCATTCGCGGAGGCTGAGGTG 176 | | |
| ORIGIN | 1963 GGAGGATCACTTGAGCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCCTGCA 2022 | | QY |
| | 175 GGAGGATCACTTGAGCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCCTGCA 116 | | |
| ORIGIN | 2023 CTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATAATGGAACA 2082 | | QY |
| | 115 CTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATAATGGAACA 56 | | |
| ORIGIN | 2083 CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATCTTTAA 2121 | | QY |
| | 55 CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATCTTTAA 17 | | |
| ORIGIN | | | Db |
| | | | |
| ORIGIN | | | RESULT 31 |
| | | | |
| ORIGIN | | | BM819720 |
| | | | |
| ORIGIN | | | LOCUS |
| | | | |
| ORIGIN | | | DEFINITION |
| | | | |
| ORIGIN | | | ACCESSION |
| | | | |
| ORIGIN | | | VERSION |
| | | | |
| ORIGIN | | | KEYWORDS |
| | | | |
| ORIGIN | | | SOURCE |
| | | | |
| ORIGIN | | | ORGANISM |
| | | | |
| ORIGIN | | | REFERENCE |
| | | | |

| | | | |
|-------------------------|--|---|-------------|
| Seq primer: M13 FORWARD | | AUTHORS | |
| POLYA=Yes. | | Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S. | |
| FEATURES | Location/Qualifiers | | TITLE |
| | 1. .552 | | |
| SOURCE | /organism="Homo sapiens" | | JOURNAL |
| | /mol_type="mRNA" | | |
| FEATURES | /db_xref="taxon:9606" | | COMMENT |
| | /clone="S18N669761-20-E04" | | |
| FEATURES | /sex="F" | | FEATURES |
| | /lab_host="Top10F" | | |
| FEATURES | /clone_lib="S18N669761" | | source |
| | /note="Organ: Stomach; Vector: pcNS; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library." | | |
| ORIGIN | 16.0%; Score 339; DB 12; Length 552; | | Query Match |
| | Best Local Similarity 100.0%; Pred. No. 6.4e-60; | | |
| ORIGIN | Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | Matches |
| | | | |
| ORIGIN | 1783 GCTCACACCTGTATCCTAGCACCTTTGGAGGCTGAGGAGGAGGATCATTGAGCCCCAG 1842 | | QY |
| | 96 GCTCACACCTGTATCCTAGCACCTTTGGAGGCTGAGGAGGAGGATCATTGAGCCCCAG 155 | | |
| ORIGIN | 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAA 1902 | | QY |
| | 156 AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAA 215 | | |
| ORIGIN | 1903 AAAATCAGCCAGTCAATGCTGGCATAACACCTGTAGTCCCAGCATTCGCGGAGGCTGAGGTG 1962 | | QY |
| | 216 AAAATCAGCCAGTCAATGCTGGCATAACACCTGTAGTCCCAGCATTCGCGGAGGCTGAGGTG 275 | | |
| ORIGIN | 1963 GGAGGATCACTTGAGCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCCTGCA 2022 | | QY |
| | 276 GGAGGATCACTTGAGCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCCTGCA 335 | | |
| ORIGIN | 2023 CTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATAATGGAACA 2082 | | QY |
| | 336 CTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATAATGGAACA 395 | | |
| ORIGIN | 2083 CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATCTTTAA 2121 | | QY |
| | 396 CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATCTTTAA 434 | | |
| ORIGIN | | | Db |
| | | | |
| ORIGIN | | | RESULT 32 |
| | | | |
| ORIGIN | | | BM819706 |
| | | | |
| ORIGIN | | | LOCUS |
| | | | |

DEFINITION K-EST0087863 S18N669761 Homo sapiens cDNA clone S18N669761-20-C12
5', mRNA sequence.
ACCESSION BM819706
VERSION BM819706.1 GI:19176119
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 567)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 20 row: C column: 12
High quality sequence stop: 567.
Location/Qualifiers
FEATURES
source
1..567
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S18N669761-20-C12"
/sex="F"
/lab_host="Top10F"
/clone_lib="S18N669761"
/note="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
ORIGIN
Query Match 16.0%; Score 339; DB 12; Length 567;
Best Local Similarity 100.0%; Pred. No. 6.4e-60;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1783 GCTCACACCTGTAATCCTAGCACCTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCCAG 1842
|||||
Db 111 GCTCACACCTGTAATCCTAGCACCTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCCAG 170
QY 1843 AAGTTGAGACTAGCTGGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAA 1902
|||||
Db 171 AAGTTGAGACTAGCTGGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAA 230
QY 1903 AAAATCAGCCAGTCATGGTGGCATAACCTGTAGTCCCAGCAATCCGGAGGCTGAGGTG 1962
|||||
Db 231 AAAATCAGCCAGTCATGGTGGCATAACCTGTAGTCCCAGCAATCCGGAGGCTGAGGTG 290
QY 1963 GGAGGATCACTTGAGCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCA 2022
|||||
Db 291 GGAGGATCACTTGAGCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCA 350
QY 2023 CTCCAGCCAGGTGACATAGCAGATCCTGTCTAAAAAATAAAAAATAAATAATGGAACA 2082
|||||

Db 351 CTCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAATAAAAAATAATATGGAACA 410
QY 2083 CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATTCTTTAA 2121
|||||
Db 411 CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATTCTTTAA 449
BM794650 570 bp mRNA linear EST 05-MAR-2002
K-EST0076037 S21SNU520 Homo sapiens cDNA clone S21SNU520-24-H09 5',
mRNA sequence.
BM794650
VERSION BM794650.1 GI:19142882
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 570)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 24 row: H column: 09
High quality sequence stop: 570.
Location/Qualifiers
FEATURES
source
1..570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNU520-24-H09"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNU520"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
ORIGIN
Query Match 16.0%; Score 338.8; DB 12; Length 570;
Best Local Similarity 81.4%; Pred. No. 7e-60;
Matches 464; Conservative 0; Mismatches 2; Indels 104; Gaps 2;
QY 1626 ATAGTTAAAAATAAATAATGTTTTAGTAAAAATGATACACTATCTCTGTGAAATAGCCTCAC 1685
|||||
Db 1 ATAGTTAAAAATAAATAATGTTTTAGTAAAAATGATACACTATCTCTGTGAAATAGCCTCAC 60
QY 1686 CCCTACATGTGGATAGAGGAATGAAAAATAATTGCTTTGACATGTCTATATGGTAC 1745

Db 61 CCCTACATGTGGATAGAGGAAATGAAAAATAATTGCTTTGACATTGTCATATGGTAC 120
QY 1746 TTTGTAAGTCATGCTTTAAGTACAAATTCATGAAA----- 1782
Db 121 TTTGTAAGTCATGCTTTAAGTACAAATTCATGAAAAGCTCACTGATCCTAATTCTTTCC 180
QY 1783 ----- 1782
Db 181 CTTTGAGGTCTCTATGCTCTGATTGTATGTATAGTAAGTGTAAAGCTGTAAAAAGTA 240
QY 1783 -----GCTCACACCTGTAATCCTAGCACCTTTGGGAGGCTGAGGAG 1822
Db 241 AATAATGTCTGGGCACAGTGGCTCAGCCTGTAAATCCTAGCACCTTTGGGAGGCTGAGGAG 300
QY 1823 GAAGGATCACTTGAGCCCGAAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTC 1882
Db 301 GAAGGATCACTTGAGCCCGAAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTC 360
QY 1883 TCTACAAAATACAGAGAGAAAAATCAGCCAGTCATGCTGGGCATACACCTGTAGTCCCAG 1942
Db 361 TCTACAAAATACAGAGAGAAAAATCAGCCAGTCATGCTGGGCCTACACCTGTAGTCCCAG 420
QY 1943 CATTCGGGAGGCTGA-GGTGGGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTG 2001
Db 421 CATTCGGGAGGCTGAGGGTGGGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTG 480
QY 2002 AGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAA 2061
Db 481 AGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAA 540
QY 2062 TAAAAAATAAATAATGGAACACAGCAAGTC 2091
Db 541 TAAAAAATAAATAATGGAACACAGCAAGTC 570

RESULT 34
BX108271/c
LOCUS BX108271 NCI_CGAP_Gas4 458 bp mRNA linear EST 06-FEB-2003
DEFINITION BX108271 IMAGE:2459693, mRNA sequence.
ACCESSION BX108271 GI:278335103
VERSION BX108271
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
JOURNAL Radelof,U., Schneider,D. and Korn,B.
COMMENT Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGp998M066117.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13u, Primer sequence: CGTTGTAAAAACGCGCCAGT.
FEATURES
source
1. .458
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGp998M066117 ; IMAGE:2459693"
/tissue_type="poorly differentiated adenocarcinoma with
signet ring cell features"
/lab_host="DH10B"
/clone_lib="NCI CGAP Gas4"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: Sali;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.69 kb. Life Technologies catalog #:
11549-011"

ORIGIN
Query Match 15.9%; Score 337.4; DB 13; Length 458;
Best Local Similarity 99.7%; Pred. No. 1.4e-59;
Matches 338; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1783 GCTCACACCTGTAATCCTAGCACCTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAG 1842
Db 351 GCTCACACCTGTAATCCTAGCACCTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAG 292
QY 1843 AAGTTCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAA 1902
Db 291 AAGTTCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAA 232
QY 1903 AAAATCAGCCAGTCATGCTGGGCATACACCTGTAGTCCCAGCATTCGCGGAGGCTGAGGTG 1962
Db 231 AAAATCAGCCGCTCATGCTGGGCATACACCTGTAGTCCCAGCATTCGCGGAGGCTGAGGTG 172
QY 1963 GGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCAGTCA 2022
Db 171 GGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCAGTCA 112
QY 2023 CTCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAATAAATAAATAATGGAACA 2082
Db 111 CTCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAATAAATAAATAATGGAACA 52
QY 2083 CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATCTTTAA 2121
Db 51 CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATCTTTAA 13

RESULT 35
AA961052/c
LOCUS AA961052 395 bp mRNA linear EST 07-JUL-1998
DEFINITION OP35b05.s1 Soares_NEL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:1578801 3', similar to contains Alu repetitive element;; mRNA
sequence.
ACCESSION AA961052
VERSION AA961052.1 GI:3127606
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 395)
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 595 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 372.
FEATURES
source
1. .395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1578801"
/lab_host="DH10B"

/clone_lib="Soares_NFL_T_GBC_S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI_CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

| | | | | |
|-----------------------|--|--------------------|-----------|-------------|
| Query Match | 15.8%; | Score 336; | DB 9; | Length 395; |
| Best Local Similarity | 100.0%; | Pred. No. 2.8e-59; | | |
| Matches 336; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY 1783 | GCTCACACCTGTAATCCTAGCACTTTGGAGGCTGAGGAGGAGGATC | ACTTGAGCCCAG | 1842 | |
| Db | | | | |
| 336 | GCTCACACCTGTAATCCTAGCACTTTGGAGGCTGAGGAGGAGGATC | ACTTGAGCCCAG | 277 | |
| QY 1843 | AAGTTCGAGACTAGCCTGGSCAACATGGAGAAGCCCTGTCTCTACA | AAAAATACAGAGAA | 1902 | |
| Db | | | | |
| 276 | AAGTTCGAGACTAGCCTGGSCAACATGGAGAAGCCCTGTCTCTACA | AAAAATACAGAGAA | 217 | |
| QY 1903 | AAATCAGCCAGTCATGTTGGCATAACACCTGTAGTCCCAGCATTC | CGGGAGGCTGAGGTG | 1962 | |
| Db | | | | |
| 216 | AAATCAGCCAGTCATGTTGGCATAACACCTGTAGTCCCAGCATTC | CGGGAGGCTGAGGTG | 157 | |
| QY 1963 | GGAGGATCACTTGAGCCCAGGAGGTTGGGGCTGCGAGTGCATGAT | CACACCACTGCA | 2022 | |
| Db | | | | |
| 156 | GGAGGATCACTTGAGCCCAGGAGGTTGGGGCTGCGAGTGCATGAT | CACACCACTGCA | 97 | |
| QY 2023 | CTCCAGCCAGGTGACATAGCGAGTCCTGTCTAAAAAATAAAAAA | ATAAATGGAACA | 2082 | |
| Db | | | | |
| 96 | CTCCAGCCAGGTGACATAGCGAGTCCTGTCTAAAAAATAAAAAA | ATAAATGGAACA | 37 | |
| QY 2083 | CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATCTT | 2118 | | |
| Db | | | | |
| 36 | CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATCTT | 1 | | |

RESULT 36

| | | | | | |
|------------|---|--------|------|--------|-----------------|
| BM785703 | BM785703 | 585 bp | mRNA | linear | EST 05-MAR-2002 |
| LOCUS | K-EST0064196 S9SNU601 Homo sapiens cDNA clone S9SNU601-32-E11 5', | | | | |
| DEFINITION | mRNA sequence. | | | | |
| ACCESSION | BM785703 | | | | |
| VERSION | BM785703.1 GI:19133935 | | | | |
| KEYWORDS | EST. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| REFERENCE | 1 (bases 1 to 585) | | | | |
| AUTHORS | Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S. | | | | |
| TITLE | 21C Frontier Korean EST Project 2001 | | | | |
| JOURNAL | Unpublished (2002) | | | | |
| COMMENT | Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsung@mail.kribb.re.kr Plate: 32 row: E column: 11 High quality sequence stop: 585. Location/Qualifiers | | | | |
| FEATURES | 1. .585 | | | | |
| Source | | | | | |

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S9SNU601-32-E11"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/cell_line="SNU-601"
/lab_host="Top10F,"
/clone_lib="S9SNU601"
/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested pME18S-FL3 vector. The
obtained cDNA vectors were used for transporation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

| | | | | |
|-----------------------|--|--------------------|-----------|-------------|
| Query Match | 15.8%; | Score 335.8; | DB 12; | Length 585; |
| Best Local Similarity | 99.4%; | Pred. No. 2.9e-59; | | |
| Matches 337; | Conservative 0; | Mismatches 2; | Indels 0; | Gaps 0; |
| QY 1783 | GCTCACACCTGTAATCCTAGCACTTTGGAGGCTGAGGAGGAGGATC | ACTTGAGCCCAG | 1842 | |
| Db | | | | |
| 137 | GCTCACGCTGTAATCCTAGCACTTTGGAGGCTGAGGAGGAGGATC | ACTTGAGCCCAG | 196 | |
| QY 1843 | AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACA | AAAAATACAGAGAA | 1902 | |
| Db | | | | |
| 197 | AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACA | AAAAATACAGAGAA | 256 | |
| QY 1903 | AAATCAGCCAGTCATGTTGGCATAACACCTGTAGTCCCAGCATTC | CGGGAGGCTGAGGTG | 1962 | |
| Db | | | | |
| 257 | AAATCAGCCAGTCATGTTGGCTACACCTGTAGTCCCAGCATTC | CGGGAGGCTGAGGTG | 316 | |
| QY 1963 | GGAGGATCACTTGAGCCCAGGAGGTTGGGGCTGCGAGTGCATGAT | CACACCACTGCA | 2022 | |
| Db | | | | |
| 317 | GGAGGATCACTTGAGCCCAGGAGGTTGGGGCTGCGAGTGCATGAT | CACACCACTGCA | 376 | |
| QY 2023 | CTCCAGCCAGGTGACATAGCGAGTCCTGTCTAAAAAATAAAAA | ATAAATGGAACA | 2082 | |
| Db | | | | |
| 377 | CTCCAGCCAGGTGACATAGCGAGTCCTGTCTAAAAAATAAAAA | ATAAATGGAACA | 436 | |
| QY 2083 | CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATCTTTAA | 2121 | | |
| Db | | | | |
| 437 | CAGCAAGTCCTAGGAAGTAGGTTAAAACTAATCTTTAA | 475 | | |

RESULT 37

| | | | | | |
|------------|---|--------|------|--------|-----------------|
| BM791792 | BM791792 | 661 bp | mRNA | linear | EST 05-MAR-2002 |
| LOCUS | K-EST0071870 S21SNU520 Homo sapiens cDNA clone S21SNU520-21-E06 5', | | | | |
| DEFINITION | mRNA sequence. | | | | |
| ACCESSION | BM791792 | | | | |
| VERSION | BM791792.1 GI:19140024 | | | | |
| KEYWORDS | EST. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| REFERENCE | 1 (bases 1 to 661) | | | | |
| AUTHORS | Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S. | | | | |
| TITLE | 21C Frontier Korean EST Project 2001 | | | | |

JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 21 row: E column: 06
High quality sequence stop: 661.
Location/Qualifiers
1. .661
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNU520-21-E06"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNU520"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 15.8%; Score 335.8; DB 12; Length 661;
Best Local Similarity 99.4%; Pred. No. 2.9e-59;
Matches 337; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1783 GCTCACACCTGTAATCCTAGCACTTTGGAGGCTGAGGAGGAAGGATCACTTGAGCCCCAG 1842
|||||
Db 215 GCTCAGCCTGTAATCCTAGCACTTTGGAGGCTGAGGAGGAAGGATCACTTGAGCCCCAG 274
|||||

Qy 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCTGTCTCTACAAAATACAGAGAGAA 1902
|||||

Db 275 AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCTGTCTCTACAAAATACAGAGAGAA 334
|||||

Qy 1903 AAAATCAGCCAGTCATGTTGGCATAACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTG 1962
|||||

Db 335 AAAATCAGCCAGTCATGTTGGCCTACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTG 394
|||||

Qy 1963 GGAGGATCACTTGAGCCAGGGAGGTTGGGCTGCAGTGAGCCATGATCACACCACTGCA 2022
|||||

Db 395 GGAGGATCACTTGAGCCAGGGAGGTTGGGCTGCAGTGAGCCATGATCACCACTGCA 454
|||||

Qy 2023 CTCAGCCAGGTGACATAGGAGATCCTGTCTAAAAAATAAATAAATAATGAACA 2082
|||||

Db 455 CTCAGCCAGGTGACATAGGAGATCCTGTCTAAAAAATAAATAAATAATGAACA 514
|||||

Qy 2083 CAGCAAGTCCTAGGAAGTAGTTAAACTAATTCITTA 2121
|||||

Db 515 CAGCAAGTCCTAGGAAGTAGTTAAACTAATTCITTA 553
|||||

RESULT 38
BB625263
LOCUS BB625263 633 bp mRNA linear EST 26-OCT-2001
DEFINITION BB625263 RIKEN full-length enriched, adult male cecum Mus musculus

cdna clone 9130209K06 5', mRNA sequence.
BB625263
VERSION BB625263.1 GI:16463563
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 633)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
TITLE Unpublished (2001)
JOURNAL
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp/
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.
Location/Qualifiers
1. .633
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="9130209K06"
/sex="male"
/tissue_type="cecum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male cecum"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5',
GAGAGAGAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse

RESULT 42
AW027675/c
LOCUS
DEFINITION
AW027675
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

AW027675 367 bp mRNA linear EST 27-OCT-1999
wv60d06.x1 NCI_CGAP Gas4 Homo sapiens cDNA clone IMAGE:2533931 3'
similar to contains_Alu repetitive element;; mRNA sequence.

AW027675
AW027675.1 GI:5886431
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 367)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. 367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2533931"
/tissue_type="poorly differentiated adenocarcinoma with
signet ring cell features"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Gas4"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.69 kb. Life Technologies catalog #:
11549-011"

ORIGIN

Query Match 15.7%; Score 333.8; DB 9; Length 367;
Best Local Similarity 99.4%; Pred. No. 8e-59;
Matches 335; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1783 GCTCACACCTGTAATCCTAGCACTTTGGAGGCTGAGGAGGAAGGATCATTGAGCCCGAG 1842
|||||
Db 337 GCTCAGCCTGTAATCCTAGCACTTTCGGAGGCTGAGGAGGAAGGATCATTGAGCCCGAG 278
|||||

QY 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAA 1902
|||||

Db 277 AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAA 218
|||||

QY 1903 AAAATCAGCCAGTCATGTTGGCATAACACCTGTAGTCCAGCATTCGCGGAGGCTGAGGTG 1962
|||||

Db 217 AAAATCAGCCAGTCATGTTGGCATAACACCTGTAGTCCAGCATTCGCGGAGGCTGAGGTG 158
|||||

QY 1963 GGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCCTGCA 2022
|||||

Db 157 GGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCCTGCA 98
|||||

QY 2023 CTCCAGCCAGGTGACATAGCGAGATCCTGTCTATAAAAAATAAAAAATAAATGAACA 2082
|||||

Db 97 CTCCAGCCAGGTGACATAGCGAGATCCTGTCTATAAAAAATAAAAAATAAATGAACA 38
|||||

QY 2083 CAGCAAGTCTAGGAAGTAGGTTAAACTAATCTTT 2119
|||||

Db 37 CAGCAAGTCTAGGAAGTAGGTTAAACTAATCTTT 1
|||||

RESULT 43
BB529462
LOCUS
DEFINITION
BB529462
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

BB529462 649 bp mRNA linear EST 26-OCT-2001
BB529462 RIKEN full-length enriched, 0 day neonate lung Mus
musculus cDNA clone E030002G01 3' similar to AF221068 Mus musculus
Claudin-18 mRNA, mRNA sequence.

BB529462
BB529462.2 GI:16445609
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 649)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
On Jul 28, 2000 this sequence version replaced gi:9580920.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:<http://genome.gsc.riken.go.jp/>
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Location/Qualifiers
1. 649
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="E030002G01"
/tissue_type="lung"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 0 day neonate"

Db324CCAGCCATGCTGCAAGCTGTACGAGCCCTGATGATCGTGGGCAATTGTTCTGGGGTCAATC383

QY366GGCCTCCTGGTATCCATCTTTGCCCCTGAAATGCATCCGCATTTGGCAGCATGGAGGACTCT425

Db384GGTATCCTCGTGTCCATCTTGCCTTGAAGTGCATTCGCATTTGGTAGCATGGATGACTCT443

QY426GCCAAGCCCAACATGACACTGACCTCCGGGATCATGTTTCATTTGTCTCAGGTCTTTGTGCA485

Db444GCCAAGGCCAAGATGACTCTGACTTCTGGGATCTTGTTCATCATCTCCGSCATCTGTGCA503

QY486ATTGCTGGAGTGTCTGTGTTTGCCCAACATGCTGGTGAACATTC530

Db504ATCATTTGGTGTCTGTGTTTGCCCAACATGCTGGTGACCAACTTC548

RESULT 45

CE210717

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CE210717

tigr-gss-dog-17000372715318

DNA

linear

GSS 25-SEP-2003

genomic survey sequence.

CE210717

CE210717.1

GI:35366372

GSS.

Canis familiaris (dog)

Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

1 (bases 1 to 731)

Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.

The dog genome: survey sequencing and comparative analysis

Science 301 (5641), 1898-1903 (2003)

22875432

14512627

Contact: Kirkness EF

The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-0200

Fax: 301-838-0208

Email: ekirknes@tigr.org

Class: shotgun.

Location/Qualifiers

1. .731

/organism="Canis familiaris"

/mol_type="genomic DNA"

/strain="Standard Poodle"

/db_xref="taxon:9615"

/clone_lib="Dog Library"

/note="Site 1: BstXI; Libraries were prepared from peripheral blood"

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches 536; Conservative

14.5%;

Score 308;

DB 29;

Length 731;

72.2%;

Pred. No. 1.7e-53;

0;

Mismatches 145;

Indels 61;

Gaps 8;

QY733GCCACAGTGTGCTACAGCCTGGAGCCTTAAGGCCAGCACTGGCTTTGGGTCCAACA792

Db20GGCCCCAATGTGCGCTTACAAGCCAGGAGGCTTCAAGGCCAGCACTGGCTTTGGGCCAGCA79

QY793CCAAAACAGAAGATATACGATGGAGGTGCCCGCACAGAGGAGGATCAATCTTATC852

Db80GCAAAAACAGAAGATATACGATGGGGTGGCCCGCACAGAGGAGGATGCAATCTCATC139

QY853CTTCCAAGCACGACTATGTGTAATGCTCTAAGACCTCTCAGCAGGGCGGAAGAACTCC912

Db140CCTCCAAGTACGACTATGTGTAGCATTTCCAAGGCACCTCAGCATGGGCAGAGAACCCT199

QY913C--GGAGAGCTCACCCAAAAAACAAGAGATCCCATCTAGATTTCTTTGCTTTTACT970

Db200CAATTAGAGCTCACTTGAAAAACAAGGAATTTTACCTCGACTCTTATTGCTTTTACT259

QY971CACAGCTGGAAGTTAGAAAAAGCCTCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCC1030

Db260CACAGTTGAAAGTAAGAAAAACCTTGATTTTCATCCGGAAGGAGGCCAGATGGTCTGGGCC319

QY1031TCAGTCTC--TGTCTCTAAATATTCCACCATAAAAAACAGCTGAGTTAT---TTATGAATTA1085

Db320ACTGCTCTGTGTTTCTAATGTTCCATCATCAAAAAACAGCTGAGTTATCAGATATGAATTA379

QY1086GAGGCTATAGCTCACATTTTCAA---TCCTCTATTTCTTTTAAATATAAACTTTCTAC1142

Db380GATGCTACAGTTCAAGTTTCCAATCCTCTGGTTTTTTTTTTTAAATTAACCTTTCTAA439

QY1143TCTGATGAGAGAATGTGTTTTTAATCTCTCTCTCACATTTTGTGATGATTAGACAGACTCC1202

Db440TCTGATGATAAAATGTGGTTTGAATTCCTATCAGACAAAT-----478

QY1203CCCTCTTCTCCTAGTCAATAAAACCCATTGATGATCTATTTCCCAGCTTATCCCCAAGAA1262

Db479-----AGCTCATCGATGATCTATTTCCCAACTCATCCCCCAGAA517

QY1263AACTTTTGAAGGAA-----AGAGTAGACCCCAAGATGTTATTTTCTGCTGTTTGAATT1316

Db518CAATTCTGAAAGTAAAGGAGAGAGTCCATCAAAAGACATCATTTCTCTACTATTTGATTT577

QY1317TTGCTCTCCCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAAGAACAAATAAGA1376

Db578TAACCTCCCCA--GCCTACTTGGCTACTAATAAACACTGATTGAAGTAG-AGCTGATAGG634

QY1377GAAAGATATTTGTAATCTCTCCAGCCCCCATGATCTCGGTTTTTCTTACACTGTGATCTTAAA1436

Db635GAAAGATATTTGTAATGCTCTCCAGCTCATTTATCTGAGTTTTTCTTACACTGGGATCTTGAC694

QY1437AGTTACCAAAACCAAGTCAATT1458

Db695ATTACCTGGACTGAAGCGATT716

Search completed: September 2, 2004, 17:25:50

Job time : 3744 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2004, 16:47:01 ; Search time 49 Seconds
(without alignments)
1675.795 Million cell updates/sec

Title: US-10-063-731-118
Perfect score: 1357
Sequence: 1 MSTTTCQVAVFLLSILGLAG.....DGGARTEDEVQSYPSKHDYV 261

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1297172 seqs, 314612898 residues

Total number of hits satisfying chosen parameters: 1297172

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------------|
| 1 | 1357 | 100.0 | 261 | 10 | US-09-946-374-326 |
| 2 | 1357 | 100.0 | 261 | 12 | US-10-206-915-356 |
| 3 | 1357 | 100.0 | 261 | 12 | US-10-199-670-356 |
| 4 | 1357 | 100.0 | 261 | 12 | US-10-201-858-356 |
| 5 | 1357 | 100.0 | 261 | 12 | US-10-205-890-356 |
| 6 | 1357 | 100.0 | 261 | 12 | US-10-208-024-356 |
| 7 | 1357 | 100.0 | 261 | 12 | US-10-201-853-356 |
| 8 | 1357 | 100.0 | 261 | 12 | US-10-063-745-118 |
| 9 | 1357 | 100.0 | 261 | 12 | US-09-940-227-76 |
| 10 | 1357 | 100.0 | 261 | 12 | US-10-063-512-118 |
| 11 | 1357 | 100.0 | 261 | 12 | US-10-063-513-118 |
| 12 | 1357 | 100.0 | 261 | 12 | US-10-063-515-118 |
| 13 | 1357 | 100.0 | 261 | 12 | US-10-063-549-118 |
| 14 | 1357 | 100.0 | 261 | 12 | US-10-063-569-118 |
| 15 | 1357 | 100.0 | 261 | 12 | US-10-063-551-118 |

| | | | | | | |
|----|------|-------|-----|----|--------------------|-------------------|
| 16 | 1357 | 100.0 | 261 | 12 | US-10-174-581-356 | Sequence 356, App |
| 17 | 1357 | 100.0 | 261 | 12 | US-10-176-483-356 | Sequence 356, App |
| 18 | 1357 | 100.0 | 261 | 12 | US-10-176-749-356 | Sequence 356, App |
| 19 | 1357 | 100.0 | 261 | 12 | US-10-176-914-356 | Sequence 356, App |
| 20 | 1357 | 100.0 | 261 | 12 | US-10-176-915-356 | Sequence 356, App |
| 21 | 1357 | 100.0 | 261 | 12 | US-10-006-485A-326 | Sequence 326, App |
| 22 | 1357 | 100.0 | 261 | 12 | US-10-013-907A-326 | Sequence 326, App |
| 23 | 1357 | 100.0 | 261 | 12 | US-10-015-499A-326 | Sequence 326, App |
| 24 | 1357 | 100.0 | 261 | 12 | US-10-063-555-118 | Sequence 118, App |
| 25 | 1357 | 100.0 | 261 | 12 | US-10-063-563-118 | Sequence 118, App |
| 26 | 1357 | 100.0 | 261 | 12 | US-10-063-594-118 | Sequence 118, App |
| 27 | 1357 | 100.0 | 261 | 12 | US-10-063-553-118 | Sequence 118, App |
| 28 | 1357 | 100.0 | 261 | 12 | US-10-063-554-118 | Sequence 118, App |
| 29 | 1357 | 100.0 | 261 | 12 | US-10-176-484-356 | Sequence 356, App |
| 30 | 1357 | 100.0 | 261 | 12 | US-10-180-550-356 | Sequence 356, App |
| 31 | 1357 | 100.0 | 261 | 12 | US-10-183-014-356 | Sequence 356, App |
| 32 | 1357 | 100.0 | 261 | 12 | US-10-187-738-356 | Sequence 356, App |
| 33 | 1357 | 100.0 | 261 | 12 | US-10-187-740-356 | Sequence 356, App |
| 34 | 1357 | 100.0 | 261 | 12 | US-10-187-883-356 | Sequence 356, App |
| 35 | 1357 | 100.0 | 261 | 12 | US-10-194-363-356 | Sequence 356, App |
| 36 | 1357 | 100.0 | 261 | 12 | US-10-194-460-356 | Sequence 356, App |
| 37 | 1357 | 100.0 | 261 | 12 | US-10-194-463-356 | Sequence 356, App |
| 38 | 1357 | 100.0 | 261 | 12 | US-10-194-484-356 | Sequence 356, App |
| 39 | 1357 | 100.0 | 261 | 12 | US-10-195-884-356 | Sequence 356, App |
| 40 | 1357 | 100.0 | 261 | 12 | US-10-195-896-356 | Sequence 356, App |
| 41 | 1357 | 100.0 | 261 | 12 | US-10-196-744-356 | Sequence 356, App |
| 42 | 1357 | 100.0 | 261 | 12 | US-10-196-755-356 | Sequence 356, App |
| 43 | 1357 | 100.0 | 261 | 12 | US-10-196-757-356 | Sequence 356, App |
| 44 | 1357 | 100.0 | 261 | 12 | US-10-197-704-356 | Sequence 356, App |
| 45 | 1357 | 100.0 | 261 | 12 | US-10-197-710-356 | Sequence 356, App |

ALIGNMENTS

RESULT 1
US-09-946-374-326
; Sequence 326, Application US/09946374
; Publication No. US20030073129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C1
; CURRENT APPLICATION NUMBER: US/09/946,374
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750

Fri Sep 3 10:15:45 2004

Fri Sep 3 10:15:45 2004

; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: 60/104987
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105000
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105002
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105104
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 60/105169
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105266
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105693
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105694
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105807

Query Match 100.0%; Score 1357; DB 10; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAVALLSILGLAGCIAATGMDMWSTQDLYDNPTVSFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVAVALLSILGLAGCIAATGMDMWSTQDLYDNPTVSFQYEGLRSCVRQSSGF 60
QY 61 TCRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TCRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANNMTGMGMVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANNMTGMGMVQTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 2
US-10-206-915-356
; Sequence 356, Application US/10206915
; Publication No. US20040029221A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C513
; CURRENT APPLICATION NUMBER: US/10/206,915
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 356
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-206-915-356

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAVALLSILGLAGCIAATGMDMWSTQDLYDNPTVSFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVAVALLSILGLAGCIAATGMDMWSTQDLYDNPTVSFQYEGLRSCVRQSSGF 60
QY 61 TCRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TCRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANNMTGMGMVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANNMTGMGMVQTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 3
US-10-199-670-356
; Sequence 356, Application US/10199670
; Publication No. US20040033560A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C401
; CURRENT APPLICATION NUMBER: US/10/199,670
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 356
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-199-670-356

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAFLLSILGLAGCIAATGMDMNSTQDLYDNPTVSFQYEGGLWRSVCVRQSSGF 60
Db 1 MSTTTCQVAFLLSILGLAGCIAATGMDMNSTQDLYDNPTVSFQYEGGLWRSVCVRQSSGF 60
QY 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 4
US-10-201-858-356
; Sequence 356, Application US/10201858
; Publication No. US2004003837A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C464
; CURRENT APPLICATION NUMBER: US/10/201,858
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 356
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-201-858-356

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAFLLSILGLAGCIAATGMDMNSTQDLYDNPTVSFQYEGGLWRSVCVRQSSGF 60
Db 1 MSTTTCQVAFLLSILGLAGCIAATGMDMNSTQDLYDNPTVSFQYEGGLWRSVCVRQSSGF 60
QY 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 5
US-10-205-890-356
; Sequence 356, Application US/10205890
; Publication No. US20040048334A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C519
; CURRENT APPLICATION NUMBER: US/10/205,890
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266

```
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 356
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-205-890-356
```

```
Query Match      100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAVALLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVAVALLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60

QY 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120

QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180

QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFSNTKNKKI 240

QY 241 YDGGARTEDEVQSYPSKHDYV 261
Db 241 YDGGARTEDEVQSYPSKHDYV 261
```

```
RESULT 6
US-10-208-024-356
; Sequence 356, Application Us/10208024
; Publication No. US20040048335A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C538
; CURRENT APPLICATION NUMBER: US/10/208,024
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
```

```
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 356
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-208-024-356

Query Match      100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAVALLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVAVALLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60

QY 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120

QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180

QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFSNTKNKKI 240

QY 241 YDGGARTEDEVQSYPSKHDYV 261
Db 241 YDGGARTEDEVQSYPSKHDYV 261
```

```
RESULT 7
US-10-201-853-356
; Sequence 356, Application US/10201853
; Publication No. US20040053358A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C465
; CURRENT APPLICATION NUMBER: US/10/201,853
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
```

```
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 356
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-201-853-356

Query Match      100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60

QY 61 TECRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120

QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMWSTANMYTGMGMVQTVQTRYTFFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMWSTANMYTGMGMVQTVQTRYTFFGAALFVGWV 180

QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKI 240

QY 241 YDGGARTEDEVQSYPSKHDYV 261
Db 241 YDGGARTEDEVQSYPSKHDYV 261

RESULT 8
US-10-063-745-118
; Sequence 118, Application US/10063745
; Publication No. US20040058411A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,745
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
```

```
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-745-118

Query Match      100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60

QY 61 TECRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120

QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMWSTANMYTGMGMVQTVQTRYTFFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMWSTANMYTGMGMVQTVQTRYTFFGAALFVGWV 180

QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKI 240

QY 241 YDGGARTEDEVQSYPSKHDYV 261
Db 241 YDGGARTEDEVQSYPSKHDYV 261

RESULT 9
US-09-940-227-76
; Sequence 76, Application US/09940227
; Publication No. US20030017468A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Sei Yu
; APPLICANT: Macina, Roberto
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific
; TITLE OF INVENTION: Genes
; FILE REFERENCE: DEX-0230
; CURRENT APPLICATION NUMBER: US/09/940,227
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,378
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 76
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-940-227-76

Query Match      100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60

QY 61 TECRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120

QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMWSTANMYTGMGMVQTVQTRYTFFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMWSTANMYTGMGMVQTVQTRYTFFGAALFVGWV 180

QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
```


QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 10

US-10-063-512-118
; Sequence 118, Application US/10063512
; Publication No. US20030018183A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,512
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-512-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPNVTSVFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPNVTSVFQYEGLRSCVRQSSGF 60
QY 61 TECRPYFTILGLPALMQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPALMQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 11

US-10-063-513-118
; Sequence 118, Application US/10063513
; Publication No. US20030018172A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,512
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-513-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,513
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-513-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPNVTSVFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPNVTSVFQYEGLRSCVRQSSGF 60
QY 61 TECRPYFTILGLPALMQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPALMQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 12

US-10-063-515-118
; Sequence 118, Application US/10063515
; Publication No. US20030018173A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,515
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-515-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPNVTSVFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPNVTSVFQYEGLRSCVRQSSGF 60
QY 61 TECRPYFTILGLPALMQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120

Db 61 TCRPYFTILGLPAMQLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSYPSKHDYV 261
Db 241 YDGGARTEDEVQSYPSKHDYV 261

RESULT 13

US-10-063-549-118
; Sequence 118, Application US/10063549
; Publication No. US20030027986A1
; GENERAL INFORMATION:

; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,549
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-549-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60
QY 61 TCRPYFTILGLPAMQLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TCRPYFTILGLPAMQLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSYPSKHDYV 261
Db 241 YDGGARTEDEVQSYPSKHDYV 261

RESULT 14

US-10-063-569-118
; Sequence 118, Application US/10063569
; Publication No. US20030018168A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,569
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-569-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60
Db 1 MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGF 60
QY 61 TCRPYFTILGLPAMQLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TCRPYFTILGLPAMQLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCAIAGSVFANMLVTNFMSTANMYTGMGMVQTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSYPSKHDYV 261
Db 241 YDGGARTEDEVQSYPSKHDYV 261

RESULT 15

US-10-063-551-118
; Sequence 118, Application US/10063551
; Publication No. US20020183494A1
; GENERAL INFORMATION:

; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,551
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-551-118

| | | | | | |
|-----------------------|------|--|---------------------|------------|-------------|
| Query Match | | 100.0%; | Score 1357; | DB 12; | Length 261; |
| Best Local Similarity | | 100.0%; | Pred. No. 2.7e-132; | | |
| Matches | 261; | Conservative | 0; | Mismatches | 0; |
| | | | | Indels | 0; |
| | | | | Gaps | 0; |
| QY | 1 | MSTTTQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGWLWSCVRQSSGF | 60 | | |
| Db | 1 | MSTTTQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGWLWSCVRQSSGF | 60 | | |
| QY | 61 | TECRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT | 120 | | |
| Db | 61 | TECRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT | 120 | | |
| QY | 121 | SGIMEFIVSGLCAIAGVSVFANMLVTNFMSTANMYTGMGGMVQTVQTRYTFGAALFVGWV | 180 | | |
| Db | 121 | SGIMEFIVSGLCAIAGVSVFANMLVTNFMSTANMYTGMGGMVQTVQTRYTFGAALFVGWV | 180 | | |
| QY | 181 | AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFSNTKNKKI | 240 | | |
| Db | 181 | AGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFSNTKNKKI | 240 | | |
| QY | 241 | YDGGARTEDEVQSPSKHDYV | 261 | | |
| Db | 241 | YDGGARTEDEVQSPSKHDYV | 261 | | |

Search completed: September 1, 2004, 16:49:39
Job time : 50 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 13:08:28 ; Search time 5592 Seconds
(without alignments)
16439.659 Million cell updates/sec

Title: US-10-063-731-117
Perfect score: 2121
Sequence: 1 gagctccctcaggagcg.....ggttaaaactaattcttta 2121

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|--------|---------------|--------|----|------------|--------------------|
| 1 | 2121 | 100.0 | 2121 | 6 | AX092386 | AX092386 Sequence |
| 2 | 2121 | 100.0 | 2121 | 6 | AX376288 | AX376288 Sequence |
| 3 | 2121 | 100.0 | 2121 | 6 | AX697257 | AX697257 Sequence |
| 4 | 2121 | 100.0 | 2121 | 9 | AY358479 | AY358479 Homo sapi |
| 5 | 1814.8 | 85.6 | 3344 | 9 | AK098474 | AK098474 Homo sapi |
| 6 | 1759 | 82.9 | 1869 | 6 | AX472954 | AX472954 Sequence |
| 7 | 1667 | 78.6 | 2040 | 9 | BC063845 | BC063845 Homo sapi |
| C 8 | 1303.2 | 61.4 | 164953 | 9 | AC016252 | AC016252 Homo sapi |
| C 9 | 1303.2 | 61.4 | 182726 | 2 | AC084436 | AC084436 Homo sapi |
| 10 | 1083.2 | 51.1 | 145477 | 2 | AC010810 | AC010810 Homo sapi |
| 11 | 954.2 | 45.0 | 982 | 6 | BD237992 | BD237992 Gastric p |
| C 12 | 845.4 | 39.9 | 1171 | 9 | HSM800576 | AL080075 Homo sapi |
| 13 | 837.2 | 39.5 | 2786 | 10 | AF221068 | AF221068 Mus muscu |
| 14 | 816 | 38.5 | 816 | 9 | AF221069 | AF221069 Homo sapi |
| 15 | 733.4 | 34.6 | 1409 | 10 | AF349450 | AF349450 Mus muscu |
| 16 | 682.8 | 32.2 | 796 | 9 | AF349452 | AF349452 Homo sapi |
| 17 | 559.4 | 26.4 | 783 | 6 | BD237993 | BD237993 Gastric p |
| 18 | 548.6 | 25.9 | 849 | 10 | AF349451 | AF349451 Mus muscu |
| 19 | 526.6 | 24.8 | 861 | 10 | AF349453 | AF349453 Mus muscu |
| C 20 | 377.8 | 17.8 | 261333 | 2 | AC138280 | AC138280 Mus muscu |
| 21 | 343.4 | 16.2 | 2388 | 5 | BC060425 | BC060425 Xenopus l |
| 22 | 220.4 | 10.4 | 423 | 6 | AX380609 | AX380609 Sequence |
| 23 | 220.4 | 10.4 | 423 | 6 | AX380725 | AX380725 Sequence |
| C 24 | 193.2 | 9.1 | 87637 | 9 | AL390960 | AL390960 Human DNA |
| 25 | 193.2 | 9.1 | 201817 | 2 | AL358992 | AL358992 Homo sapi |
| 26 | 191 | 9.0 | 129402 | 9 | AC011510 | AC011510 Homo sapi |
| C 27 | 190.2 | 9.0 | 166229 | 2 | AC022070 | AC022070 Homo sapi |
| 28 | 190 | 9.0 | 4297 | 9 | HSM802170 | AL137457 Homo sapi |
| 29 | 189.8 | 8.9 | 168672 | 2 | AC027102 | AC027102 Homo sapi |
| C 30 | 189.8 | 8.9 | 168958 | 9 | AC027451 | AC027451 Homo sapi |
| C 31 | 189.8 | 8.9 | 183842 | 9 | AC130450 | AC130450 Homo sapi |
| C 32 | 189.8 | 8.9 | 188636 | 9 | HUAC002287 | AC002287 Homo sapi |
| C 33 | 189.8 | 8.9 | 200863 | 9 | AC018608 | AC018608 Homo sapi |
| 34 | 189.4 | 8.9 | 150262 | 9 | AC034186 | AC034186 Homo sapi |
| 35 | 189.4 | 8.9 | 167579 | 9 | AC026161 | AC026161 Homo sapi |
| C 36 | 189.2 | 8.9 | 125823 | 9 | HSJ561L24 | AL049796 Human DNA |
| C 37 | 189 | 8.9 | 128379 | 9 | HSJ773A18 | AL049557 Human DNA |
| C 38 | 189 | 8.9 | 161333 | 2 | AC141071 | AC141071 Homo sapi |
| 39 | 189 | 8.9 | 181298 | 2 | AC142540 | AC142540 Homo sapi |
| C 40 | 189 | 8.9 | 189155 | 2 | AC143324 | AC143324 Homo sapi |
| C 41 | 189 | 8.9 | 189363 | 9 | AC126603 | AC126603 Homo sapi |
| C 42 | 189 | 8.9 | 189582 | 2 | AC142534 | AC142534 Homo sapi |
| 43 | 189 | 8.9 | 196587 | 2 | AC141449 | AC141449 Homo sapi |
| 44 | 189 | 8.9 | 200389 | 9 | AC134678 | AC134678 Homo sapi |
| 45 | 189 | 8.9 | 203928 | 2 | AC145311 | AC145311 Homo sapi |

ALIGNMENTS

RESULT 1
AX092386
LOCUS AX092386
DEFINITION Sequence 117 from Patent WO0116318.
ACCESSION AX092386
VERSION AX092386.1 GI:13444507
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A.,
Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and
Wood,W.I.

AX092386 2121 bp DNA Linear PAT 21-MAR-2001
Sequence 117 from Patent WO0116318.

| | | | | | | | | | | |
|----------|---|--|---------------|-----------|--------------|--|--|--|--|--|
| TITLE | Secreted and transmembrane polypeptides and nucleic acids encoding the same | | | | | | | | | |
| JOURNAL | Patent: WO 0116318-A 117 08-MAR-2001; Genentech, Inc. (US) | | | | | | | | | |
| FEATURES | Location/Qualifiers | | | | | | | | | |
| source | 1..2121 | | | | | | | | | |
| | /organism="Homo sapiens" | | | | | | | | | |
| | /mol_type="unassigned DNA" | | | | | | | | | |
| | /db_xref="taxon:9606" | | | | | | | | | |
| ORIGIN | | | | | | | | | | |
| | Query Match | 100.0%; | Score 2121; | DB 6; | Length 2121; | | | | | |
| | Best Local Similarity | 100.0%; | Pred. No. 0; | | | | | | | |
| | Matches 2121; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; | | | | | |
| QY | 1 | GAGCTCCCTCAGGAGCGGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCT | 60 | | | | | | | |
| Db | 1 | GAGCTCCCTCAGGAGCGGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCT | 60 | | | | | | | |
| QY | 61 | CGCAGGCGGAGGGCGGCGCCAGGATCATGTCCACCACCAACATGTCACCAAGTGGCGT | 120 | | | | | | | |
| Db | 61 | CGCAGGCGGAGGGCGGCGCCAGGATCATGTCCACCACCAACATGTCACCAAGTGGCGT | 120 | | | | | | | |
| QY | 121 | TCCTCCTGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGATGGACATGTGGA | 180 | | | | | | | |
| Db | 121 | TCCTCCTGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGATGGACATGTGGA | 180 | | | | | | | |
| QY | 181 | GCACCCAGGACCTGTACGACAAACCCCGTCACCTCCGTGTTCCAGTACGAAGGCTCTGGA | 240 | | | | | | | |
| Db | 181 | GCACCCAGGACCTGTACGACAAACCCCGTCACCTCCGTGTTCCAGTACGAAGGCTCTGGA | 240 | | | | | | | |
| QY | 241 | GGAGCTGCGTGAGGCAGAGTTTCAGGCTTCACCGAATGAGGCCCTATTTACCATCCTGG | 300 | | | | | | | |
| Db | 241 | GGAGCTGCGTGAGGCAGAGTTTCAGGCTTCACCGAATGAGGCCCTATTTACCATCCTGG | 300 | | | | | | | |
| QY | 301 | GACTTCCAGCCATGCTGCAGGAGTGCAGGCCCTGATGATCGTAGGCATCGTCCTGGTG | 360 | | | | | | | |
| Db | 301 | GACTTCCAGCCATGCTGCAGGAGTGCAGGCCCTGATGATCGTAGGCATCGTCCTGGTG | 360 | | | | | | | |
| QY | 361 | CCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTTGGCAGCATGGAG | 420 | | | | | | | |
| Db | 361 | CCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTTGGCAGCATGGAG | 420 | | | | | | | |
| QY | 421 | ACTCTGCCAAAGCCAAACATGACACTGACCTCCGGGATCATGTTTCTCAGGCTCTTT | 480 | | | | | | | |
| Db | 421 | ACTCTGCCAAAGCCAAACATGACACTGACCTCCGGGATCATGTTTCTCAGGCTCTTT | 480 | | | | | | | |
| QY | 481 | GTGCAATTGCTGGAGTGTCTGTGTTTGCCAAACATGCTGGTGAATACTTCTGGATGTCCA | 540 | | | | | | | |
| Db | 481 | GTGCAATTGCTGGAGTGTCTGTGTTTGCCAAACATGCTGGTGAATACTTCTGGATGTCCA | 540 | | | | | | | |
| QY | 541 | CAGCTAACATGTATACACCGGCATGGTGGGATGGTGACAGTGTTCAGACCTGTTT | 600 | | | | | | | |
| Db | 541 | CAGCTAACATGTATACACCGGCATGGTGGGATGGTGACAGTGTTCAGACCTGTTT | 600 | | | | | | | |
| QY | 601 | TTGGTGGGCTCTGTTCTGGGCTGGGTCGGTGGAGGCCCTCACACTAATTGGGGGTGA | 660 | | | | | | | |
| Db | 601 | TTGGTGGGCTCTGTTCTGGGCTGGGTCGGTGGAGGCCCTCACACTAATTGGGGGTGA | 660 | | | | | | | |
| QY | 661 | TGATGTGCATCGCTGCCGGGCTGGCACCAGAAACCAACTACAAAGCCGTTTCTT | 720 | | | | | | | |
| Db | 661 | TGATGTGCATCGCTGCCGGGCTGGCACCAGAAACCAACTACAAAGCCGTTTCTT | 720 | | | | | | | |
| QY | 721 | ATCATGCCTCAGGCCACAGTGTTCCTTACAGCCCTGGAGGCTTCAAGGCCAGCACTGGCT | 780 | | | | | | | |
| Db | 721 | ATCATGCCTCAGGCCACAGTGTTCCTTACAGCCCTGGAGGCTTCAAGGCCAGCACTGGCT | 780 | | | | | | | |
| QY | 781 | TTGGGTCCAAACACCAAAACCAAGAGATATACGATGGAGGTGCCCGCACAGAGGACGAG | 840 | | | | | | | |
| Db | 781 | TTGGGTCCAAACACCAAAACCAAGAGATATACGATGGAGGTGCCCGCACAGAGGACGAG | 840 | | | | | | | |
| QY | 841 | TACAATCTTATCCTTCCAAAGCAGCATATGTGTAATGCTCTAAGACCTCTCAGCACGGGC | 900 | | | | | | | |

| | | | |
|----|------|--|------|
| Db | 841 | TACAATCTTATCCTTCCAAGCAGCAGCTATGTGTAATGCTCTAAGACCTCTCAGCACGGC | 900 |
| QY | 901 | GGAAAGAACTCCCGGAGAGCTCACCCCAAAAAACAAGGATCCCATCTAGATTCTTCTT | 960 |
| Db | 901 | GGAAAGAACTCCCGGAGAGCTCACCCCAAAAAACAAGGATCCCATCTAGATTCTTCTT | 960 |
| QY | 961 | GCTTTTGTACTCAGCTGGAAGTTAGAAAAGCCTCGATTTCATCTTTGGAGAGGCCAAAT | 1020 |
| Db | 961 | GCTTTTGTACTCAGCTGGAAGTTAGAAAAGCCTCGATTTCATCTTTGGAGAGGCCAAAT | 1020 |
| QY | 1021 | GGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACAGCTGAGTTATTATG | 1080 |
| Db | 1021 | GGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACAGCTGAGTTATTATG | 1080 |
| QY | 1081 | AATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTAAATATAACTTTCT | 1140 |
| Db | 1081 | AATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTAAATATAACTTTCT | 1140 |
| QY | 1141 | ACTCTGATGAGAGAAATGTGGTTTAAATCTCTCTCTCTCTCAATTTTGATTTAGACAGCT | 1200 |
| Db | 1141 | ACTCTGATGAGAGAAATGTGGTTTAAATCTCTCTCTCTCTCAATTTTGATTTAGACAGCT | 1200 |
| QY | 1201 | CCCCCTCTTCTCTAGTCAATAAACCCCAATGATGATCTATTTCCAGCTTATCCCCAAG | 1260 |
| Db | 1201 | CCCCCTCTTCTCTAGTCAATAAACCCCAATGATGATCTATTTCCAGCTTATCCCCAAG | 1260 |
| QY | 1261 | AAAACTTTTGAAGGAAAGAGTAGACCCCAAGATGTTATTTCTGCTGTTTGAATTTGT | 1320 |
| Db | 1261 | AAAACTTTTGAAGGAAAGAGTAGACCCCAAGATGTTATTTCTGCTGTTTGAATTTGT | 1320 |
| QY | 1321 | CTCCCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAAGCAATAAGAGAAA | 1380 |
| Db | 1321 | CTCCCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAAGCAATAAGAGAAA | 1380 |
| QY | 1381 | GATATTGTAATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATCTTAAAGTT | 1440 |
| Db | 1381 | GATATTGTAATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATCTTAAAGTT | 1440 |
| QY | 1441 | ACCAAAACCAAGTCATTTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTGTTGACATCT | 1500 |
| Db | 1441 | ACCAAAACCAAGTCATTTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTGTTGACATCT | 1500 |
| QY | 1501 | TCATTATACAGCAACACCATTTTAGGAGTTTCTCAGCTCTCCACTGGAGTCTCTTCT | 1560 |
| Db | 1501 | TCATTATACAGCAACACCATTTTAGGAGTTTCTCAGCTCTCCACTGGAGTCTCTTCT | 1560 |
| QY | 1561 | GTGCGGGTCAGAAATTGTCCTTAGATGAATGAGAAAATATTTTAAATTAAGTCC | 1620 |
| Db | 1561 | GTGCGGGTCAGAAATTGTCCTTAGATGAATGAGAAAATATTTTAAATTAAGTCC | 1620 |
| QY | 1621 | TAAATATAGTTAAATAAATATGTTTAGTAAATGATACACTATCTCTGTGAATAGC | 1680 |
| Db | 1621 | TAAATATAGTTAAATAAATAATGTTTAGTAAATGATACACTATCTCTGTGAATAGC | 1680 |
| QY | 1681 | CTCACCCCTACATGTGGATAGAGGAAATGAAAAATAATGCTTTGACATTGCTCTATAT | 1740 |
| Db | 1681 | CTCACCCCTACATGTGGATAGAGGAAATGAAAAATAATGCTTTGACATTGCTCTATAT | 1740 |
| QY | 1741 | GGTACTTTGTAAAGTCATGCTTAACTAATAAATTCATGAAAAGCTCACACCTGTAATCCT | 1800 |
| Db | 1741 | GGTACTTTGTAAAGTCATGCTTAACTAATAAATTCATGAAAAGCTCACACCTGTAATCCT | 1800 |
| QY | 1801 | AGCACTTTGGAGGCTGAGGAGGAGGATCACTTGAGCCCAAGATTCGAGACTAGCCTG | 1860 |
| Db | 1801 | AGCACTTTGGAGGCTGAGGAGGAGGATCACTTGAGCCCAAGATTCGAGACTAGCCTG | 1860 |
| QY | 1861 | GGCAACATGGAGAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCATGG | 1920 |
| Db | 1861 | GGCAACATGGAGAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCATGG | 1920 |
| QY | 1921 | TGGCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCC | 1980 |
| Db | 1921 | TGGCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCC | 1980 |

QY 1981 AGGAGGTTGGGCTGCAGTGAGCCATGATCACACCACCTGCACCTCCAGCCAGGTGACATA 2040
|||||
Db 1981 AGGAGGTTGGGCTGCAGTGAGCCATGATCACACCACCTGCACCTCCAGCCAGGTGACATA 2040
|||||
QY 2041 GCGAGATCCTGTCTAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCTCTAGGAAGT 2100
|||||
Db 2041 GCGAGATCCTGTCTAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCTCTAGGAAGT 2100
|||||
QY 2101 AGGTTAAAACTAATTCCTTAA 2121
|||||
Db 2101 AGGTTAAAACTAATTCCTTAA 2121
|||||

RESULT 2
AX376288
LOCUS AX376288 2121 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 355 from Patent WO0168848.
ACCESSION AX376288
VERSION AX376288.1 GI:19170535
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Baker,K.P., Chen,J., Desnoyers,L., Goddard,A., Godowski,P.J.,
Gurney,A.L., Pan,J., Smith,V., Watanabe,C.K., Wood,W.I. and
Zhang,Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0168848-A 355 20-SEP-2001;
Genentech, Inc. (US)
FEATURES
source Location/Qualifiers
1..2121
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2121; DB 6; Length 2121;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTCCCCCTCAGGAGCGGTTAGCTTCACACCTCGGCAGCAGGAGCGCGCAGCTTCT 60
|||||
Db 1 GAGTCCCCCTCAGGAGCGGTTAGCTTCACACCTCGGCAGCAGGAGCGCGCAGCTTCT 60
|||||

QY 61 GCGAGGCGCAGGGCGGCGGCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGT 120
|||||
Db 61 GCGAGGCGCAGGGCGGCGGCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGT 120
|||||

QY 121 TCCTCCTGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGA 180
|||||
Db 121 TCCTCCTGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGA 180
|||||

QY 181 GCACCCAGGACCTGTACGACAAACCCCGTCACCTCCGTTCAGTACGAAGGGCTCTGGA 240
|||||
Db 181 GCACCCAGGACCTGTACGACAAACCCCGTCACCTCCGTTCAGTACGAAGGGCTCTGGA 240
|||||

QY 241 GGAGCTGCGTGAGGCAGAGTTCAGGTTTACCGAATGCAGGCCCTATTTTACCATCCTGG 300
|||||
Db 241 GGAGCTGCGTGAGGCAGAGTTTACGGTTTACCGAATGCAGGCCCTATTTTACCATCCTGG 300
|||||

QY 301 GACTTCCAGCATGCTGCAGGCAGTGCAGCCCTGATGATCGTAGGCATCGTCTGGGTG 360
|||||
Db 301 GACTTCCAGCATGCTGCAGGCAGTGCAGCCCTGATGATCGTAGGCATCGTCTGGGTG 360
|||||

QY 361 CCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTCGACATGGAGG 420
|||||
Db 361 CCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTCGACATGGAGG 420
|||||

QY 421 ACTCTGCCAAGCCAAACATGACACTGACCTCCGGGATCATGTTCTCAGGTCTTT 480
|||||

Db 421 ACTCTGCCAAGCCAAACATGACACTGACCTCCGGATCATGTTCTCAGGTCTTT 480
|||||
QY 481 GTGCAATTGCTGGAGTGTCTGTGTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCA 540
|||||
Db 481 GTGCAATTGCTGGAGTGTCTGTGTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCA 540
|||||
QY 541 CAGCTAACATGTACACCGGCATGGTGGGATGGTGCAGACTGTTCAGACCAGGTACACAT 600
|||||
Db 541 CAGCTAACATGTACACCGGCATGGTGGGATGGTGCAGACTGTTCAGACCAGGTACACAT 600
|||||
QY 601 TTGGTGGGCTCTGTTCTGTTGGGTGGGTGGTGGAGGCTCACACTAATTGGGGGTGTA 660
|||||
Db 601 TTGGTGGGCTCTGTTCTGTTGGGTGGGTGGTGGAGGCTCACACTAATTGGGGGTGTA 660
|||||
QY 661 TGATGTGCATCGCCTGCCGGGCTGGCACACAGAAACCAACTACAAAGCCGTTTCTT 720
|||||
Db 661 TGATGTGCATCGCCTGCCGGGCTGGCACACAGAAACCAACTACAAAGCCGTTTCTT 720
|||||
QY 721 ATCATGCCTCAGGCCACAGTGTTCCTACAAGCCTGGAGGCTTCAAGCCAGCACTGGCT 780
|||||
Db 721 ATCATGCCTCAGGCCACAGTGTTCCTACAAGCCTGGAGGCTTCAAGCCAGCACTGGCT 780
|||||
QY 781 TTGGTCCAACACCAACCAAGAGATATACGATGAGGTGCCCGCACAGAGCAGG 840
|||||
Db 781 TTGGTCCAACACCAACCAAGAGATATACGATGAGGTGCCCGCACAGAGCAGG 840
|||||
QY 841 TACAATCTTATCCTTCCAAGCAGACTATGTGTAATGCTCTAAGACCTCTCAGCACGGGC 900
|||||
Db 841 TACAATCTTATCCTTCCAAGCAGACTATGTGTAATGCTCTAAGACCTCTCAGCACGGGC 900
|||||
QY 901 GGAAGAAACTCCCGGAGACTCACCCAAAACAAAGGAGATCCCATCTAGATTTCTTCT 960
|||||
Db 901 GGAAGAAACTCCCGGAGACTCACCCAAAACAAAGGAGATCCCATCTAGATTTCTTCT 960
|||||
QY 961 GCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTTCACTCTTTGGAGAGGCCAAAT 1020
|||||
Db 961 GCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTTCACTCTTTGGAGAGGCCAAAT 1020
|||||
QY 1021 GGTCTTAGCCTCAGTCTCTCTCTAAATATTCACCATAAAACAGCTGAGTTATTTATG 1080
|||||
Db 1021 GGTCTTAGCCTCAGTCTCTCTCTAAATATTCACCATAAAACAGCTGAGTTATTTATG 1080
|||||
QY 1081 AATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTAAATATAACTTTCT 1140
|||||
Db 1081 AATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTAAATATAACTTTCT 1140
|||||
QY 1141 ACTCTGATGAGAGATGTGGTTTTAATCTCTCTCTCACATTTTGAATTTAGACAGCT 1200
|||||
Db 1141 ACTCTGATGAGAGATGTGGTTTTAATCTCTCTCTCACATTTTGAATTTAGACAGCT 1200
|||||
QY 1201 CCCCCTCTTCTCTCTAGTCAATAAACCCATGATGATCTATTTCCAGCTTATCCCCAAG 1260
|||||
Db 1201 CCCCCTCTTCTCTCTAGTCAATAAACCCATGATGATCTATTTCCAGCTTATCCCCAAG 1260
|||||
QY 1261 AAAAATTTTGAAGAAAGAGTAGACCCAAAGATGTTATTTCTGCTGTTTGAATTTGT 1320
|||||
Db 1261 AAAAATTTTGAAGAAAGAGTAGACCCAAAGATGTTATTTCTGCTGTTTGAATTTGT 1320
|||||
QY 1321 CTCCCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAAGCAATAAGAGAAA 1380
|||||
Db 1321 CTCCCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAAGCAATAAGAGAAA 1380
|||||
QY 1381 GATAATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTTCTTACTGTGATCTTAAAGTT 1440
|||||
Db 1381 GATAATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTTCTTACTGTGATCTTAAAGTT 1440
|||||
QY 1441 ACCAAACCAAAGTCATTTTTCAGTTTGAGGCAACCAACCTTCTACTGCTGTGACATCT 1500
|||||
Db 1441 ACCAAACCAAAGTCATTTTTCAGTTTGAGGCAACCAACCTTCTACTGCTGTGACATCT 1500
|||||
QY 1501 TCTTATTACAGCAACACCATTTCTAGGAGTTTCTGAGTCTCCACTGGAGTCTCTTTCT 1560
|||||

Db 1501 TCTTATTACAGCAACACCATTTCTAGGAGTTTCTGAGCTTCACACTGCGAGTCCCTCTTTCT 1560
QY 1561 GTCGGGGTTCAGAAATGTCCCTAGATGAATGAGAAAATTTTAAATTTAAAGTCC 1620
Db 1561 GTCGGGGTTCAGAAATGTCCCTAGATGAATGAGAAAATTTTAAATTTAAAGTCC 1620
QY 1621 TAAATATAGTTAAATAAATAATGTTTCTAGTAAATGATACACTATCTCTGTGAAATAGC 1680
Db 1621 TAAATATAGTTAAATAAATAATGTTTCTAGTAAATGATACACTATCTCTGTGAAATAGC 1680
QY 1681 CTCACCCCTACATGTGGATAGAGGAATGAAAAAATAATGCTTTGACATTTGCTATAT 1740
Db 1681 CTCACCCCTACATGTGGATAGAGGAATGAAAAAATAATGCTTTGACATTTGCTATAT 1740
QY 1741 GGTACTTTGTAAAGTCACTGTTAAGTACAAATCCATGAAAGCTCACACCTGTAATCCT 1800
Db 1741 GGTACTTTGTAAAGTCACTGTTAAGTACAAATCCATGAAAGCTCACACCTGTAATCCT 1800
QY 1801 AGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAAGTTCGAGACTAGCCTG 1860
Db 1801 AGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAAGTTCGAGACTAGCCTG 1860
QY 1861 GCGAATGAGAGAGCCCTGTCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCATGG 1920
Db 1861 GCGAATGAGAGAGCCCTGTCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCATGG 1920
QY 1921 TGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTGGAGGATCACTTGAGCCC 1980
Db 1921 TGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTGGAGGATCACTTGAGCCC 1980
QY 1981 AGGAGGTTGGGCTGAGTGAAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATA 2040
Db 1981 AGGAGGTTGGGCTGAGTGAAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATA 2040
QY 2041 GCGAGATCCTGTCTAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCTTAGGAAGT 2100
Db 2041 GCGAGATCCTGTCTAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCTTAGGAAGT 2100
QY 2101 AGGTTAAAACTAATCTTTAA 2121
Db 2101 AGGTTAAAACTAATCTTTAA 2121

RESULT 3
AX697257 2121 bp DNA linear PAT 02-APR-2003
LOCUS AX697257 Sequence 325 from Patent WO0078961.
DEFINITION AX697257
ACCESSION AX697257
VERSION AX697257.1 GI:29498417
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ferrara,N., Stewart,T.A., Williams,P.M., Baker,K.P., Desnoyers,L.,
Eaton,D.L., Gao,W.Q., Pan,J., Botstein,D., Fong,S., Goddard,A.,
Godowski,P.J., Gurney,A.L., Smith,V., Tumas,D., Wood,W.I.,
Grimaldi,C.J., Hillan,K.J., Paoni,N.F., Roy,M.A. and Watanabe,C.K.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0078961-A 325 28-DEC-2000;
Genentech Inc. (US)
FEATURES
source 1..2121
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2121; DB 6; Length 2121;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTCCCTCAGGAGCGGTTAGCTTACACCTTCGGCAGCAGGAGGGCGGCGCTTCT 60
Db 1 GAGTCCCTCAGGAGCGGTTAGCTTACACCTTCGGCAGCAGGAGGGCGGCGCTTCT 60
QY 61 CGCAGGCGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 61 CGCAGGCGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 121 TCCTCCTGTCCATCCTGGGGCTGGCGGCTGCATCGCGGCGGCGGCGGCGGCGGCGG 180
Db 121 TCCTCCTGTCCATCCTGGGGCTGGCGGCTGCATCGCGGCGGCGGCGGCGGCGGCGG 180
QY 181 GCACCCAGGACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGCTCTGGA 240
Db 181 GCACCCAGGACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGCTCTGGA 240
QY 241 GGAGCTGCGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 241 GGAGCTGCGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 301 GACTTCCAGCCATGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Db 301 GACTTCCAGCCATGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 361 CCATTTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCAATGGCAGCATGGAGG 420
Db 361 CCATTTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCAATGGCAGCATGGAGG 420
QY 421 ACTCTGCCAAAGCCAAACATGACACTGACCTCCGGGATCATGTTCAATGCTCAGGTCTTT 480
Db 421 ACTCTGCCAAAGCCAAACATGACACTGACCTCCGGGATCATGTTCAATGCTCAGGTCTTT 480
QY 481 GTGCAATTTGCTGGAGTGTCTGTTTGGCAACATGCTGGTGAATACTTTGGATGTCCA 540
Db 481 GTGCAATTTGCTGGAGTGTCTGTTTGGCAACATGCTGGTGAATACTTTGGATGTCCA 540
QY 541 CAGCTAACATGTACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Db 541 CAGCTAACATGTACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
QY 601 TTGTTGGGCTCTGTTTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 660
Db 601 TTGTTGGGCTCTGTTTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 660
QY 661 TGATGTGCATCGCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCT 720
Db 661 TGATGTGCATCGCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCT 720
QY 721 ATCATGCTCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
Db 721 ATCATGCTCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
QY 781 TTGGGTCACCAACACCAAAACAGAGATATACGATGGAGGTGCCCGCACAGAGGACGAGG 840
Db 781 TTGGGTCACCAACACCAAAACAGAGATATACGATGGAGGTGCCCGCACAGAGGACGAGG 840
QY 841 TACAATCTTATCCTTCCAAGCAGCTATGTGTAATGCTCTAAGACCTCTCAGCAGCGGC 900
Db 841 TACAATCTTATCCTTCCAAGCAGCTATGTGTAATGCTCTAAGACCTCTCAGCAGCGGC 900
QY 901 GGAAGAAACTCCCGGAGAGCTCACCCAAAAAACAAGGAGATCCCATCTAGATTTCTTCTT 960
Db 901 GGAAGAAACTCCCGGAGAGCTCACCCAAAAAACAAGGAGATCCCATCTAGATTTCTTCTT 960
QY 961 GCTTTTGAATCAGAGTGAAGTTAGAAAAGCCTCGATTTCATCTTTGGAGAGGCCAAAT 1020
Db 961 GCTTTTGAATCAGAGTGAAGTTAGAAAAGCCTCGATTTCATCTTTGGAGAGGCCAAAT 1020
QY 1021 GGTCTTAGCCTCAGTCTCTGTCTCTAATAATTTCCACCATAAACAGCTGAGTTATTATG 1080
Db 1021 GGTCTTAGCCTCAGTCTCTGTCTCTAATAATTTCCACCATAAACAGCTGAGTTATTATG 1080

QY 1081 AATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTAAATATAACTTTCT 1140
Db |||||
QY 1081 AATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTAAATATAACTTTCT 1140
Db |||||
QY 1141 ACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAGACT 1200
Db |||||
QY 1141 ACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAGACT 1200
QY 1201 CCCCTCTTCTCTCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAG 1260
Db |||||
QY 1201 CCCCTCTTCTCTCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAG 1260
QY 1261 AAAACTTTTGAAGGAAAGAGTAGACCCCAAAGATGTTATTTTCTGCTGTTTGAATTTGT 1320
Db |||||
QY 1261 AAAACTTTTGAAGGAAAGAGTAGACCCCAAAGATGTTATTTTCTGCTGTTTGAATTTGT 1320
QY 1321 CTCGCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAAGAAGCAATAAGAGAAA 1380
Db |||||
QY 1321 CTCGCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAAGAAGCAATAAGAGAAA 1380
QY 1381 GATATTTGTAATCTCTCCAGGCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAAAGTT 1440
Db |||||
QY 1381 GATATTTGTAATCTCTCCAGGCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAAAGTT 1440
QY 1441 ACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTGTGACATCT 1500
Db |||||
QY 1441 ACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTGTGACATCT 1500
QY 1501 TCTTATTACAGCAACACCAATTTAGGAGTTTCTCTGAGCTCTCCACTGGAGTCCTTTCT 1560
Db |||||
QY 1501 TCTTATTACAGCAACACCAATTTAGGAGTTTCTCTGAGCTCTCCACTGGAGTCCTTTCT 1560
QY 1561 GTCGGGGTGAGAAATTTGCCCTAGATGAATGAGAAAAATATTTTTTTAAATTAAGTCC 1620
Db |||||
QY 1561 GTCGGGGTGAGAAATTTGCCCTAGATGAATGAGAAAAATATTTTTTTAAATTAAGTCC 1620
QY 1621 TAAATATAGTTAAATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGC 1680
Db |||||
QY 1621 TAAATATAGTTAAATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGC 1680
QY 1681 CTCACCCCTACATGTGGATAGAAAGGAAATGAAAAAATAATGCTTTTGACATTTGCTATAT 1740
Db |||||
QY 1681 CTCACCCCTACATGTGGATAGAAAGGAAATGAAAAAATAATGCTTTTGACATTTGCTATAT 1740
QY 1741 GGTACTTTGTAAAGTCATGCTTAAGTACAAATTCATGAAAGCTCACACCTGTAATCCT 1800
Db |||||
QY 1741 GGTACTTTGTAAAGTCATGCTTAAGTACAAATTCATGAAAGCTCACACCTGTAATCCT 1800
QY 1801 AGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAAGTTTCGAGACTAGCCCTG 1860
Db |||||
QY 1801 AGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAAGTTTCGAGACTAGCCCTG 1860
QY 1861 GGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCATGG 1920
Db |||||
QY 1861 GGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCATGG 1920
QY 1921 TGGCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGAGGATCACTTGAGCCC 1980
Db |||||
QY 1921 TGGCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGAGGATCACTTGAGCCC 1980
QY 1981 AGGAGGTTGGGCTGCAGTGAGCCATGATCACACCACCTGCACCTCCAGCCAGGTGACATA 2040
Db |||||
QY 1981 AGGAGGTTGGGCTGCAGTGAGCCATGATCACACCACCTGCACCTCCAGCCAGGTGACATA 2040
QY 2041 GCGAGATCCTGTCTAAAAAATAAAAAATAAATGGAACACAGCAAGTCCTAGGAAGT 2100
Db |||||
QY 2041 GCGAGATCCTGTCTAAAAAATAAAAAATAAATGGAACACAGCAAGTCCTAGGAAGT 2100
QY 2101 AGGTTAAAACTAATTTCTTTAA 2121
Db |||||
QY 2101 AGGTTAAAACTAATTTCTTTAA 2121

RESULT 4
AY358479
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AY358479 2121 bp mRNA linear PRI 03-OCT-2003
Homo sapiens clone DNA73734 CLDN18 (UNQ778) mRNA, complete cds.
AY358479
FLI_CDNA.
1 GI:37182080
Homo sapiens (human)
Homo sapiens

REFERENCE
AUTHORS

1 (bases 1 to 2121)
Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,
Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B.,
Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E.,
Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S.,
Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C.,
Schoenfeld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,V.,
Stinson,J., Vagts,A., Vandlen,R., Watanabe,C., Wieand,D., Woods,K.,
Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z.,
Goddard,A., Wood,W.I. and Godowski,P.
The Secreted Protein Discovery Initiative (SPDI), a Large-Scale
Effort to Identify Novel Human Secreted and Transmembrane Proteins:
A Bioinformatics Assessment
Genome Res. 13 (10), 2265-2270 (2003)
12975309
2 (bases 1 to 2121)
Clark,H.F.
Direct Submission

TITLE

Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,
Inc., 1 DNA Way, South San Francisco, CA 94080, USA

JOURNAL
PUBMED
REFERENCE
AUTHORS

Location/Qualifiers
1. .2121
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DNA73734"
1. .2121
/locus_tag="UNQ778"
90. .875
/locus_tag="UNQ778"
/note="PRO1572"
/codon_start=1
/product="CLDN18"
/protein_id="AAQ88843.1"
/db_xref="GI:37182081"
/translation="MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVF
QYEGLRSCVRQSSGFTECRPYFTILGLPAMLQAVRALMIVGLGAILGLVSIPLK
CIRIGSMEDSAKANMTLTSGIMFIVSGLCIAGVSFANMLVTNFMSTANMYTGMGG
MVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIACGLAPEETNYKAVSYHASGHSV
AYKPGGFKASTGFGSNTKNKIYDGGARTEDEVQSYPSKHDYV"

TITLE
JOURNAL

Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,
Inc., 1 DNA Way, South San Francisco, CA 94080, USA

FEATURES
source

gene

CDS

ORIGIN

Query Match 100.0%; Score 2121; DB 9; Length 2121;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTCCCTCAGGAGCGGTTAGCTTCACACCTTCGSCAGCAGAGGGCGGCGCTTCT 60
Db |||||
QY 1 GAGCTCCCTCAGGAGCGGTTAGCTTCACACCTTCGSCAGCAGAGGGCGGCGCTTCT 60
Db |||||
QY 61 CGCAGGCGGCGAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 120
Db |||||
QY 61 CGCAGGCGGCGAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 120
QY 121 TCCTCCTGTCCATCCTGGGGCTGGCGGCGGCTGCATCGCGGCCACCGGATGGACATGTGGA 180
Db |||||
QY 121 TCCTCCTGTCCATCCTGGGGCTGGCGGCGGCTGCATCGCGGCCACCGGATGGACATGTGGA 180
QY 181 GCACCCAGGACCTGTACGACAAACCCCGTCACCTCCGTTCCAGTACGAAGGGCTCTGGA 240
Db |||||
QY 181 GCACCCAGGACCTGTACGACAAACCCCGTCACCTCCGTTCCAGTACGAAGGGCTCTGGA 240

| | | | | | | | |
|----|------|---|------|--|--|---|------|
| QY | 241 | GGAGCTGCGTGGAGCAGAGTTTCAGGCTTACCGAATGCAGGCCCTATTTTCACCATCCTGG | 300 | | 1321 | CTCCCCACCCCAACTTGGCTAGTAATAACAATTACTGAAGAAGCAATAAGAGAA | 1380 |
| Db | 241 | GGAGCTGCGTGGAGCAGAGTTTCAGGCTTACCGAATGCAGGCCCTATTTTCACCATCCTGG | 300 | | 1381 | GATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAAAGTT | 1440 |
| QY | 301 | GACTTCCAGCCATGCTGCAGGCAGTGGAGCCCTGATGATCGTAGGCATCGTCTGGGTG | 360 | | 1381 | GATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAAAGTT | 1440 |
| Db | 301 | GACTTCCAGCCATGCTGCAGGCAGTGGAGCCCTGATGATCGTAGGCATCGTCTGGGTG | 360 | | 1441 | ACCAAAACCAAGTCATTTTCAGTTTGAGGGCAACCAACCTTTCTACTGCTGTTGACATCT | 1500 |
| QY | 361 | CCATTGGCCCTCTGGTATCCATCTTTGGCCCTGAAATGCATCCGCATTTGGCAGCATGGAGG | 420 | | 1441 | ACCAAAACCAAGTCATTTTCAGTTTGAGGGCAACCAACCTTTCTACTGCTGTTGACATCT | 1500 |
| Db | 361 | CCATTGGCCCTCTGGTATCCATCTTTGGCCCTGAAATGCATCCGCATTTGGCAGCATGGAGG | 420 | | 1501 | TCTTATTACAGCAACACCATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTTCT | 1560 |
| QY | 421 | ACTCTGCCAAAGCCAAACATGACACTGACCTCCGGGATCATGTTCAATTGTTCTCAGGTCCTT | 480 | | 1501 | TCTTATTACAGCAACACCATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTTCT | 1560 |
| Db | 421 | ACTCTGCCAAAGCCAAACATGACACTGACCTCCGGGATCATGTTCAATTGTTCTCAGGTCCTT | 480 | | 1561 | GTCGGGGTCAGAAATTGTCCCTAGATGAATGAGAAATTAATTTTTTTTAATTTAAGTCC | 1620 |
| QY | 481 | GTGCAATTGCTGGAGTGTCTGTGTTTGCCAAACATGCTGGTGACATACTTCTGGATGTCCA | 540 | | 1561 | GTCGGGGTCAGAAATTGTCCCTAGATGAATGAGAAATTAATTTTTTTTAATTTAAGTCC | 1620 |
| Db | 481 | GTGCAATTGCTGGAGTGTCTGTGTTTGCCAAACATGCTGGTGACATACTTCTGGATGTCCA | 540 | | 1621 | TAAATATAGTTTAAATAATAATGTTTTTAGTAAATGATACACTATCTCTGTGAAATAGC | 1680 |
| QY | 541 | CAGCTAACATGTACACCGGCATGGTGGATGGTGACACTGTTTCAGACCAAGGTACACAT | 600 | | 1621 | TAAATATAGTTTAAATAATAATGTTTTTAGTAAATGATACACTATCTCTGTGAAATAGC | 1680 |
| Db | 541 | CAGCTAACATGTACACCGGCATGGTGGATGGTGACACTGTTTCAGACCAAGGTACACAT | 600 | | 1681 | CTCACCCCTACATGTGGATAGAAAGGAATGAAAAATAATTTGCTTTGACATTTGCTATAT | 1740 |
| QY | 601 | TTGGTGGGGCTCTGTTCTGGGCTGGTCCCTACAGCCCTGGAGGCTTCAAGGCCAGCACTGGCT | 660 | | 1681 | CTCACCCCTACATGTGGATAGAAAGGAATGAAAAATAATTTGCTTTGACATTTGCTATAT | 1740 |
| Db | 601 | TTGGTGGGGCTCTGTTCTGGGCTGGTCCCTACAGCCCTGGAGGCTTCAAGGCCAGCACTGGCT | 660 | | 1741 | GCTACTTTGTAAAGTCATGCTTAAGTACAAATTCATGAAAGCTCACACCTGTAATCCT | 1800 |
| QY | 661 | TGATGTGCATCGCCTGCCGGGCTGGCACCAGAGAAACCAACTACAAAGCCGTTTCTT | 720 | | 1741 | GCTACTTTGTAAAGTCATGCTTAAGTACAAATTCATGAAAGCTCACACCTGTAATCCT | 1800 |
| Db | 661 | TGATGTGCATCGCCTGCCGGGCTGGCACCAGAGAAACCAACTACAAAGCCGTTTCTT | 720 | | 1801 | AGCACTTTGGGAGGCTGAGGAGGAGGATCACTTGAGCCAGAAAGTTTCGAGACTAGCCTG | 1860 |
| QY | 721 | ATCATGCTCAGGCCACAGTGTGCTTACAGCCCTGGAGGCTTCAAGGCCAGCACTGGCT | 780 | | 1801 | AGCACTTTGGGAGGCTGAGGAGGAGGATCACTTGAGCCAGAAAGTTTCGAGACTAGCCTG | 1860 |
| Db | 721 | ATCATGCTCAGGCCACAGTGTGCTTACAGCCCTGGAGGCTTCAAGGCCAGCACTGGCT | 780 | | 1861 | GGCAACATGGAGAACCCCTGTCTACAAATAACAGAGAGAAAAATCAGCCAGTCATGG | 1920 |
| QY | 781 | TTGGTCCAAACACCAACCAAGAAAGATATACGATGGAGGTGCCCGCACAGAGACGAGG | 840 | | 1861 | GGCAACATGGAGAACCCCTGTCTACAAATAACAGAGAGAAAAATCAGCCAGTCATGG | 1920 |
| Db | 781 | TTGGTCCAAACACCAACCAAGAAAGATATACGATGGAGGTGCCCGCACAGAGACGAGG | 840 | | 1921 | TGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTTGAGCCC | 1980 |
| QY | 841 | TACAACTTATCTTCCAGCAGCACTATGTGTAATGCTTAAGACCTCTCAGCACGGGC | 900 | | 1921 | TGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTTGAGCCC | 1980 |
| Db | 841 | TACAACTTATCTTCCAGCAGCACTATGTGTAATGCTTAAGACCTCTCAGCACGGGC | 900 | | 1981 | AGGAGGTTGGGCTGCAGTGAGCCATGATCACACCTGCACCTCCAGCCAGGTGACATA | 2040 |
| QY | 901 | GGAAGAACTCCCGGAGAGCTCACCCAAAAACAAGAGATCCCATCTAGATTCTTCTT | 960 | | 1981 | AGGAGGTTGGGCTGCAGTGAGCCATGATCACACCTGCACCTCCAGCCAGGTGACATA | 2040 |
| Db | 901 | GGAAGAACTCCCGGAGAGCTCACCCAAAAACAAGAGATCCCATCTAGATTCTTCTT | 960 | | 2041 | GCGAGATCCTGTCTAAAAATAAAAAATAATAATGGAACACAGCAAGTCTTAGGAAGT | 2100 |
| QY | 961 | GCTTTTGACTCAGCTGGAAGTTAGAAAAGCCTCGATTTCATCTTTGGAGAGGCCAAAT | 1020 | | 2041 | GCGAGATCCTGTCTAAAAATAAAAAATAATAATGGAACACAGCAAGTCTTAGGAAGT | 2100 |
| Db | 961 | GCTTTTGACTCAGCTGGAAGTTAGAAAAGCCTCGATTTCATCTTTGGAGAGGCCAAAT | 1020 | | 2101 | AGGTTAAAACTAATTTCTTTAA | 2121 |
| QY | 1021 | GGTCTTAGCCTCAGTCTGTCTCTAAATATTCACCATAAACAGCTGAGTTATTATG | 1080 | | 2101 | AGGTTAAAACTAATTTCTTTAA | 2121 |
| Db | 1021 | GGTCTTAGCCTCAGTCTGTCTCTAAATATTCACCATAAACAGCTGAGTTATTATG | 1080 | | 2101 | AGGTTAAAACTAATTTCTTTAA | 2121 |
| QY | 1081 | AATTAGAGGTATAGCTCACATTTTCAATCTCTATTTCTTTTAAATATACTTTCT | 1140 | | RESULT 5 | | |
| Db | 1081 | AATTAGAGGTATAGCTCACATTTTCAATCTCTATTTCTTTTAAATATACTTTCT | 1140 | | AK098474 | | |
| QY | 1141 | ACTCTGATGAGAGAATGTGGTTTAAATCTCTCTCTCACATTTTGGATGATTAGACAGACT | 1200 | | LOCUS | | |
| Db | 1141 | ACTCTGATGAGAGAATGTGGTTTAAATCTCTCTCTCACATTTTGGATGATTAGACAGACT | 1200 | | DEFINITION | | |
| QY | 1201 | CCCCCTCTTCTCTAGTCAATAAACCCATTTGATGATCTATTTCCAGCTTATCCCAAG | 1260 | | AK098474 | | |
| Db | 1201 | CCCCCTCTTCTCTAGTCAATAAACCCATTTGATGATCTATTTCCAGCTTATCCCAAG | 1260 | | ACCESION | | |
| QY | 1261 | AAAACTTTTGAAGGAAAGTAGACCCCAAGATGTTATTTCTGCTGTTGAATTTTGT | 1320 | | VERSION | | |
| Db | 1261 | AAAACTTTTGAAGGAAAGTAGACCCCAAGATGTTATTTCTGCTGTTGAATTTTGT | 1320 | | AK098474.1 | | |
| QY | 1321 | CTCCCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAGCAATAAGAGAA | 1380 | | KEYWORDS | | |
| Db | 1321 | CTCCCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAGCAATAAGAGAA | 1380 | | oligo capping; fis (full insert sequence). | | |
| | | | | | SOURCE | | |
| | | | | | Homo sapiens (human) | | |
| | | | | | ORGANISM | | |
| | | | | | Homo sapiens | | |
| | | | | | REFERENCE | | |
| | | | | | 1 | | |
| | | | | | Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., | | |
| | | | | | Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., | | |
| | | | | | Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., | | |
| | | | | | Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., | | |

Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
 Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and
 Sugano,S.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 3344)
 Sugano,S. and Suzuki,Y.
 Direct Submission
 Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure, Human Genome
 Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction and 5'-end one pass sequencing: Institute of Medical
 Science, University of Tokyo, Laboratory of Genome Structure, Human
 Genome Center; 3'-end one pass sequencing: RAB; clone selection for
 full insert sequencing: RAB and Helix Research Institute.
 Location/Qualifiers
 1. .3344
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="STM00359"
 /tissue_type="stomach mucosa"
 /clone_lib="STM"
 /note="cloning vector: pME18SFL3"
 source
 source
 COMMENT
 COMMENT
 FEATURES
 source
 ORIGIN
 ORIGIN

| Query Match | 85.6% | Score 1814.8 | DB 9 | Length 3344 |
|-----------------------|--------------|---|---------------|-------------|
| Best Local Similarity | 92.0% | Pred. No. 0 | | |
| Matches 1968 | Conservative | 0 | Mismatches 67 | Indels 103 |
| | | | Gaps | 1 |
| QY | 87 | ATCATGTCCACCAACACATGCCAAGTGGTGGCGTTCTCTCTGTCCATCTCTGGGCTGGCC | 146 | |
| Db | 51 | ACCATGGCCGTGACTGCCTGTCCAGGGCTTGGGTTCTGTTTCACTGATGGGATTGCG | 110 | |
| QY | 147 | GGCTGCATCGGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACAAACCCC | 206 | |
| Db | 111 | GGCATCATGTGTCACCTGCATGGACCACTGAGCAACCAAGACTTGTACAAACACCCC | 170 | |
| QY | 207 | GTCACTCCGTGTTCCAGTACGAAGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTTCAGGC | 266 | |
| Db | 171 | GTAACAGCTGTTTTCAACTACGAGGGCTGTGGCGCTCTGTGTCGAGAGAGCTCTGGC | 230 | |
| QY | 267 | TTCACCGAATGCAGGCCCTATTTCACCATCTCTGGGACTTCCAGCCATGCTGCAGGCAGTG | 326 | |
| Db | 231 | TTCACCGAGTCCCGGGGCTACTTCACCTGCTGGGCTGCCAGCCATGCTGCAGGCAGTG | 290 | |
| QY | 327 | CGAGCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTGGCCTCCTGGTATCCATCTTT | 386 | |
| Db | 291 | CGAGCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTGGCCTCCTGGTATCCATCTTT | 350 | |
| QY | 387 | GCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACACTG | 446 | |
| Db | 351 | GCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACACTG | 410 | |
| QY | 447 | ACCTCCGGGATCATGTTTCAATTGCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTT | 506 | |
| Db | 411 | ACCTCCGGGATCATGTTTCAATTGCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTT | 470 | |
| QY | 507 | GCCAAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGCATGGGT | 566 | |
| Db | 471 | GCCAAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGCATGGGT | 530 | |
| QY | 567 | GGGATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGCGTCTGTTTCGTGGGCTGG | 626 | |
| Db | 531 | GGGATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGCGTCTGTTTCGTGGGCTGG | 590 | |
| QY | 627 | GTGCTGGAGGCCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCTGCCGGGGCCTG | 686 | |

QY 1767 ACAAATTCATCAAAA----- 1782
Db 1731 ACAAATTCATCAAAAAGCTCACTGATCCTAAATCTTTCCCTTTGAGGTCTCTATGGCTCT 1790
QY 1783 -----G 1783
Db 1791 GATTGTACATGATAGTAAGTAAAGCCATGTAAAAAGTAAATAATGTCTGGGCACAGTGG 1850
QY 1784 CTCACACCTGTATCTTAGCACTTTGGGAGGCTGAGGAGGAAGATCACTTGAGCCCCAGA 1843
Db 1851 CTCACACCTGTATCTTAGCACTTTGGGAGGCTGAGGAGGAAGATCACTTGAGCCCCAGA 1910
QY 1844 AGTTCGAGACTAGCTGGGCAACATGGAGAGCCCTGTCTCTACAAAAATACAGAGAGAAA 1903
Db 1911 AGTTCGAGACTAGCTGGCAACATGGAGAGCCCTGTCTCTACAAAAATACAGAGAGAAA 1970
QY 1904 AAATCAGCCAGTCATGGTGGCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGG 1963
Db 1971 AAATCAGCCAGTCATGGTGGCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGG 2030
QY 1964 GAGGATCACTTGAGCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCAC 2023
Db 2031 GAGGATCACTTGAGCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCAC 2090
QY 2024 TCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAATAAAAAATAAATAATGGAACAC 2083
Db 2091 TCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAATAAAAAATAAATAATGGAACAC 2150
QY 2084 AGCAAGTCTAGGAAGTAGGTAAATACTAATCTTTAA 2121
Db 2151 AGCAAGTCTAGGAAGTAGGTAAATACTAATCTTTAA 2188

RESULT 6
AX472954 1869 bp DNA linear PAT 09-AUG-2002
LOCUS
DEFINITION Sequence 3 from Patent WO0218576.
ACCESSION AX472954
VERSION AX472954.1 GI:22207741
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chen, S.Y., Macina, R.A., Sun, Y. and Recipon, H.
TITLE Compositions and methods relating to lung specific genes
JOURNAL Patent: WO 0218576-A 3 07-MAR-2002;
Diadexus, Inc. (US)
FEATURES
source
1. .1869
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 82.9%; Score 1759; DB 6; Length 1869;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1781; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 6 CCCCTCAGGAGCGGTTAGCTTACACCTTCGGCAGCAGGAGGGCGGCGAGCTTCTCGCAG 65
Db 1 CCCCTCAGGAGCGGTTA-CTTCAACCTTCGGCAGCAGGAGGGCGGCA-CTTCTCGCAG 58
QY 66 GCGGCAGGGCGGCGGCGGATCATGTCCACCACCATGCCAAGTGGTGGCTTCTC 125
Db 59 GCGGCAGGGCGGCGGCGGATCATGTCCACCACCATGCCAAGTGGTGGCTTCTC 118
QY 126 CTGTCCATCCTGGGGTGGCGGCTGCATCGCGGCCACCGGATGGAATGTGGAGACC 185
Db 119 CTGTCCATCCTGGGGTGGCGGCTGCATCGCGGCCACCGGATGGAATGTGGAGACC 178

QY 186 CAGGACCTGTACGACAAACCCCGTCACTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGC 245
Db 179 CAGGACCTGTACGACAAACCCCGTCACTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGC 238
QY 246 TGCCTGAGGAGAGTTTCAAGGCTTCAAGGATGAGGAGGCTTATTTACCATCTCTGGGACTT 305
Db 239 TGCCTGAGGAGAGTTTCAAGGCTTCAAGGATGAGGAGGCTTATTTACCATCTCTGGGACTT 298
QY 306 CCAGCCATGTGTCAGGAGGAGTGGAGGCTGATGATCGTAGGATCGTCTGGTGGCTT 365
Db 299 CCAGCCATGTGTCAGGAGGAGTGGAGGCTGATGATCGTAGGATCGTCTGGTGGCTT 358
QY 366 GGCCTCCTGATFCCATCTTTGCCCTGAAATGCATCCGATTTGGCAGCATGGAGGACTCT 425
Db 359 GGCCTCCTGATFCCATCTTTGCCCTGAAATGCATCCGATTTGGCAGCATGGAGGACTCT 418
QY 426 GCCAAAGCCAAACATGACACTGACCTCCGGGATCATGTTTCAATTTGATGATCTTTGTGCA 485
Db 419 GCCAAAGCCAAACATGACACTGACCTCCGGGATCATGTTTCAATTTGATGATCTTTGTGCA 478
QY 486 ATTGCTGGAGTGTCTGTGTTTCCCAACATGCTGGTGAATACTTTGGATGTCCACAGCT 545
Db 479 ATTGCTGGAGTGTCTGTGTTTCCCAACATGCTGGTGAATACTTTGGATGTCCACAGCT 538
QY 546 AACATGTACACCGGATGGTGGATGGTGCAGACTGTTTCCAGACCAAGGTACACATTTGGT 605
Db 539 AACATGTACACCGGATGGTGGATGGTGCAGACTGTTTCCAGACCAAGGTACACATTTGGT 598
QY 606 GCGGCTCTGTTGCTGGGCTGGTGGTGGTGCAGACTGTTTGGGGTGTGATGATG 665
Db 599 GCGGCTCTGTTGCTGGGCTGGTGGTGGTGCAGACTGTTTGGGGTGTGATGATG 658
QY 666 TGCATCGCTGCGGGGCTGGCACCAGAGAGAAACCAACTACAAAGCCGTTTCTTATCAT 725
Db 659 TGCATCGCTGCGGGGCTGGCACCAGAGAGAAACCAACTACAAAGCCGTTTCTTATCAT 718
QY 726 GCTCAGGCCACAGTGTGCTTACAGCCTGAGGCTTCAAGGCCAGACTGGCTTTGGG 785
Db 719 GCTCAGGCCACAGTGTGCTTACAGCCTGAGGCTTCAAGGCCAGACTGGCTTTGGG 778
QY 786 TCCAAACACCAAAACAAAGAGATATACGATGGAGTGCCTGCACAGAGAGGAGTACAA 845
Db 779 TCCAAACACCAAAACAAAGAGATATACGATGGAGTGCCTGCACAGAGAGGAGTACAA 838
QY 846 TCTTATCTTCCAGCAGCAGTGTGTAATGCTCTAAGACCTCTCAGCACGGCGGGAAG 905
Db 839 TCTTATCTTCCAGCAGCAGTGTGTAATGCTCTAAGACCTCTCAGCACGGCGGGAAG 898
QY 906 AAACCTCCGGGAGAGCTCACCCAAACAAACAGGAGATCCCATCTAGATTTCTTCTGCTTT 965
Db 899 AAACCTCCGGGAGAGCTCACCCAAACAAACAGGAGATCCCATCTAGATTTCTTCTGCTTT 958
QY 966 TGACTCAGCTGGAGGTTAGAAAAGCCTCGATTTCTTTGGAGAGGCCAAATGGTCT 1025
Db 959 TGACTCAGCTGGAGGTTAGAAAAGCCTCGATTTCTTTGGAGAGGCCAAATGGTCT 1018
QY 1026 TAGCCTCAGTCTCTGCTCTAATAATTCACCATATAACAGAGTATTTATGAATTA 1085
Db 1019 TAGCCTCAGTCTCTGCTCTAATAATTCACCATATAACAGAGTATTTATGAATTA 1078
QY 1086 GAGGCTATAGCTCACATTTTCAATCTCTATTTCTTTTAAATATACTTTTCTACTCT 1145
Db 1079 GAGGCTATAGCTCACATTTTCAATCTCTATTTCTTTTAAATATACTTTTCTACTCT 1138
QY 1146 GATGAGAGATGTGGTTTAAATCTCTCTCACAATTTTGTGATGATTTAGACAGACTCCCC 1205
Db 1139 GATGAGAGATGTGGTTTAAATCTCTCTCACAATTTTGTGATGATTTAGACAGACTCCCC 1198
QY 1206 TCTTCTCTAGTCAATAAACCCATGATGATCTATTTCCAGCTTATCCCCAAGAAAC 1265
Db 1199 TCTTCTCTAGTCAATAAACCCATGATGATCTATTTCCAGCTTATCCCCAAGAAAC 1258
QY 1266 TTTTGAAAGGAGAGTAGACCCAAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTCCC 1325

Db 1259 TTTTGAAGAAAGAGTAGACCCAAAGATGTTATTTCTGCTGTTTGAATTTTGTCTCCC 1318
QY 1326 CACCCCAACTTGGCTAGTAATAACACTTACTGAAGAAGCAATAAGAGAAAGATAT 1385
Db 1319 CACCCCAACTTGGCTAGTAATAACACTTACTGAAGAAGCAATAAGAGAAAGATAT 1378
QY 1386 TTGTAATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATCTTAAAGTTACCAA 1445
Db 1379 TTGTAATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATCTTAAAGTTACCAA 1438
QY 1446 ACCAAAGTCATTTTCAGTTTGAGCAACCAACCTTTCTACTGCTGTTGACATCTTCTTA 1505
Db 1439 ACCAAAGTCATTTTCAGTTTGAGCAACCAACCTTTCTACTGCTGTTGACATCTTCTTA 1498
QY 1506 TTACAGCAACACCATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCCTTTCTGTGTCG 1565
Db 1499 TTACAGCAACACCATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCCTTTCTGTGTCG 1558
QY 1566 GGGTCAGAAATTGTCCTTAGATGAATGAGAAAATTATTTTAAJTTAAGTCCTAAAT 1625
Db 1559 GGGTCAGAAATTGTCCTTAGATGAATGAGAAAATTATTTTAAJTTAAGTCCTAAAT 1618
QY 1626 ATAGTTAAATAAATAATGTTTAGTAAATGATACACTATCTCTGTGAATAGCCTCAC 1685
Db 1619 ATAGTTAAATAAATAATGTTTAGTAAATGATACACTATCTCTGTGAATAGCCTCAC 1678
QY 1686 CCCTACATGTGGATAGAGGAATGAAAAATAATTGCTTTGACATTGCTATATGGTAC 1745
Db 1679 CCCTACATGTGGATAGAGGAATGAAAAATAATTGCTTTGACATTGCTATATGGTAC 1738
QY 1746 TTTGTAAGTCATGCTTTAAGTACAAATTCCATGAAAGCTCAC 1788
Db 1739 TTTGTAAGTCATGCTTTAAGTACAAATTCCATGAAAGCTCAC 1781

RESULT 7
BC063845
LOCUS BC063845
DEFINITION Homo sapiens cDNA clone IMAGE:6177081, partial cds.
ACCESSION BC063845
VERSION BC063845.1 GI:39795429
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2040)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
REFERENCE 2 (bases 1 to 2040)
AUTHORS Strausberg,R.
TITLE Direct Submission

JOURNAL
REMARK
COMMENT

Submitted (08-DEC-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 141 Row: a Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7705960.

FEATURES
source

Location/Qualifiers
1..2040
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6177081"
/tissue_type="Peripheral Nervous System, dorsal root
ganglion"
/clone_lib="Lupski_dorsal_root_ganglion"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 78.6%; Score 1667; DB 9; Length 2040;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 10; Indels 221; Gaps 2;
QY 80 GGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCTTCTCTGTCCATCCTGGG 139
Db 1 GGCCAGGATCATGTGTCCACCACCACATGCCAAGTGGTGGCTTCTCTGTCCATCCTGGG 60
QY 140 GCTGGCCGGTGCATCGCGGCCACCAGGATGGACATGTGGAGCACCCAGGACCTGTACGA 199
Db 61 GCTGGCCGGTGCATCGCGGCCACCAGGATGGACATGTGGAGCACCCAGGACCTGTACGA 120
QY 200 CAAACCCCGTCACCTCGTGTTCAGTACGAAGGGCTCTGGAGGAGCTCGTGAGGCAGAG 259
Db 121 CAAACCCCGTCACCTCGTGTTCAGTACGAAGGGCTCTGGAGGAGCTCGTGAGGCAGAG 180
QY 260 TTCAGGCTTCACCGAATGCAGGCCCTATTTTACCACATCCTGGGACTTCCAGCCATGCTGCA 319
Db 181 TTCAGGCTTCACCGAATGCAGGCCCTATTTTACCACATCCTGGGACTTCCAGCCATGCTGCA 240
QY 320 GGCAGTGGAGCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTTGGCCTCCTGGTATC 379
Db 241 GGCAGTGGAGCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTTGGCCTCCTGGTATC 300
QY 380 CATCTTTGCCCTGAAATGCATCCGATGGCAGCATGGAGACTCTGCCAAAGCCAAACAT 439
Db 301 CATCTTTGCCCTGAAATGCATCCGATGGCAGCATGGAGACTCTGCCAAAGCCAAACAT 360
QY 440 GACACTGACCTCCGGGATCATGTTCAATTTCTCAGGTCTTTGTGCAATTGCTGGAGTGC 499
Db 361 GACACTGACCTCCGGGATCATGTTCAATTTCTCAGGTCTTTGTGCAATTGCTGGAGTGC 392
QY 500 TGTGTTTGCCAAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACGG 559

Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 164953)
Worley,K.C.

Direct Submission
Submitted (24-NOV-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 164953)
Worley,K.C.

Direct Submission
Submitted (06-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 164953)
Worley,K.C.

Direct Submission
Submitted (28-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 164953)
Worley,K.C.

Direct Submission
Submitted (25-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 164953)
Worley,K.C.

Direct Submission
Submitted (25-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Mar 28, 2002 this sequence version replaced gi:19172796.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

| FEATURES | | QUALSTAT-REPORT. | |
|---------------|---------------------|---|--|
| source | Location/Qualifiers | | |
| misc_feature | 1. .164953 | /organism="Homo sapiens" | |
| | | /mol_type="genomic DNA" | |
| | | /db_xref="taxon:9606" | |
| | | /chromosome="3" | |
| | | /clone="RP11-461M2" | |
| | | complement(1. .2055) | |
| | | /note="overlaps bases 1. .2055 of clone AC023049" | |
| | | /function="clone overlap" | |
| | | complement(81. .571) | |
| | | /rpt_family="L1MC/D" | |
| repeat_region | | complement(802. .1221) | |
| repeat_region | | /rpt_family="L1M4" | |
| repeat_region | | complement(1351. .3211) | |
| repeat_region | | /rpt_family="L1MB2" | |
| repeat_region | | complement(4306. .4615) | |
| repeat_region | | /rpt_family="AluSx" | |
| repeat_region | | 5734. .5828 | |
| repeat_region | | /rpt_family="L2" | |
| repeat_region | | 5932. .6109 | |
| repeat_region | | /rpt_family="L2" | |
| repeat_region | | complement(6172. .6471) | |
| repeat_region | | /rpt_family="AluJb" | |
| repeat_region | | 6496. .7131 | |
| repeat_region | | /rpt_family="L2" | |
| repeat_region | | 6500. .7300 | |
| repeat_region | | /rpt_family="L2" | |
| repeat_region | | /function="trinucleotide tandem repeat" | |
| repeat_region | | 7144. .7401 | |
| repeat_region | | /rpt_family="L2" | |
| repeat_region | | 7509. .7591 | |
| repeat_region | | /rpt_family="L2" | |
| repeat_region | | complement(9509. .9612) | |
| repeat_region | | /rpt_family="L1MB6" | |
| repeat_region | | complement(10594. .10729) | |
| repeat_region | | /rpt_family="L1MB6" | |
| repeat_region | | 10810. .11005 | |
| repeat_region | | /rpt_family="L1MB6" | |
| repeat_region | | 11077. .11124 | |
| repeat_region | | /rpt_family="L2" | |
| repeat_region | | 11922. .11943 | |
| repeat_region | | /rpt_family="L2" | |
| repeat_region | | 12284. .12589 | |
| repeat_region | | /rpt_family="L2" | |
| repeat_region | | 12914. .13614 | |
| repeat_region | | /rpt_family="L1MC1" | |
| repeat_region | | complement(14500. .14605) | |
| repeat_region | | /rpt_family="L1MC1" | |
| repeat_region | | complement(16916. .17151) | |
| repeat_region | | /rpt_family="L1MEC" | |
| repeat_region | | complement(17156. .18175) | |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 182726)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 3, clone RP11-468E18
Unpublished
2 (bases 1 to 182726)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Barna,N., Bastien,V., Bida,F., Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Piegore,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (04-NOV-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 182726)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 15, 2001 this sequence version replaced gi:11094829.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11425
Center clone name: 468_E18
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 168531 bases at least Q40
Consensus quality: 175180 bases at least Q30

Consensus quality: 177830 bases at least Q20
Insert size: 179326; sum-of-contigs
Quality coverage: 4.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 35 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1
* 1676: contig of 1676 bp in length
* 1677: gap of 100 bp
* 1777: contig of 891 bp in length
* 2667: gap of 100 bp
* 2668: gap of 100 bp
* 2767: gap of 100 bp
* 3120: contig of 353 bp in length
* 3121: gap of 100 bp
* 3221: contig of 699 bp in length
* 3919: gap of 100 bp
* 4019: gap of 100 bp
* 4921: contig of 902 bp in length
* 4922: gap of 100 bp
* 5021: contig of 1391 bp in length
* 6412: gap of 100 bp
* 6512: gap of 100 bp
* 7808: contig of 1296 bp in length
* 7908: gap of 100 bp
* 9082: contig of 1174 bp in length
* 9182: gap of 100 bp
* 9083: gap of 100 bp
* 11518: contig of 2336 bp in length
* 11519: gap of 100 bp
* 11619: contig of 2809 bp in length
* 14527: gap of 100 bp
* 14528: gap of 100 bp
* 14528: contig of 1913 bp in length
* 16440: gap of 100 bp
* 16441: gap of 100 bp
* 16541: contig of 2800 bp in length
* 19340: gap of 100 bp
* 19341: contig of 3166 bp in length
* 22607: gap of 100 bp
* 22707: contig of 2509 bp in length
* 25215: gap of 100 bp
* 25316: contig of 2535 bp in length
* 27851: gap of 100 bp
* 27951: contig of 19970 bp in length
* 47921: gap of 100 bp
* 48021: contig of 2592 bp in length
* 50613: gap of 100 bp
* 50713: contig of 2345 bp in length
* 53057: gap of 100 bp
* 53058: contig of 1703 bp in length
* 53158: gap of 100 bp
* 54861: contig of 2085 bp in length
* 54961: gap of 100 bp
* 57046: contig of 3830 bp in length
* 57146: gap of 100 bp
* 60975: contig of 4857 bp in length
* 61075: gap of 100 bp
* 61076: contig of 4012 bp in length
* 65933: gap of 100 bp
* 66033: contig of 4012 bp in length
* 70045: gap of 100 bp
* 70145: contig of 6437 bp in length
* 76581: gap of 100 bp
* 76582: contig of 6806 bp in length
* 83488: gap of 100 bp
* 83588: contig of 5447 bp in length
* 89035: gap of 100 bp
* 89135: contig of 6276 bp in length
* 95411: gap of 100 bp
* 95511: contig of 8055 bp in length
* 103566: gap of 100 bp
* 103666: contig of 8930 bp in length
* 112595: gap of 100 bp
* 112596: contig of 9965 bp in length
* 112696: gap of 100 bp
* 122661: contig of 8991 bp in length
* 122761: gap of 100 bp

* 131752 131851: gap of 100 bp
* 131852 144507: contig of 12656 bp in length
* 144508 144607: gap of 100 bp
* 144608 160481: contig of 15874 bp in length
* 160482 160581: gap of 100 bp
* 160582 181386: contig of 20805 bp in length
* 181387 181486: gap of 100 bp
* 181487 182726: contig of 1240 bp in length.

FEATURES

source
1. .182726
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3"
/clone="RP11-468E18"
/clone_lib="RPCI-11 Human Male BAC"
1. .1676
/note="assembly_fragment"
clone_end:SP6
vector_side:left
misc_feature
1777. _2667
/note="assembly_fragment"
2768. .3120
/note="assembly_fragment"
3221. .3919
/note="assembly_fragment"
4020. .4921
/note="assembly_fragment"
5022. .6412
/note="assembly_fragment"
6513. .7808
/note="assembly_fragment"
7909. .9082
/note="assembly_fragment"
9183. .11518
/note="assembly_fragment"
11619. .14427
/note="assembly_fragment"
14528. .16440
/note="assembly_fragment"
16541. .19340
/note="assembly_fragment"
19441. .22606
/note="assembly_fragment"
22707. .25215
misc_feature

Query Match 61.4%; Score 1303.2; DB 2; Length 182726;
Best Local Similarity 93.0%; Pred. No. 6.6e-287;
Matches 1418; Conservative 0; Mismatches 3; Indels 103; Gaps 1;
QY 701 CAACCTACAAAGCCGTTCTTATCATGCCTCAGGCCACAGTGTTCCTACAAAGCCTGGAGG 760
Db 47532 CAGCTACAAAGCCGTTCTTATCATGCCTCAGGCCACAGTGTTCCTACAAAGCCTGGAGG 47473
QY 761 CTTCAAGCCAGCACTGGCTTTGGGTCCAACACCAAAACAAGAAGATATACGATGGAGG 820
Db 47472 CTTCAAGCCAGCACTGGCTTTGGGTCCAACACCAAAACAAGAAGATATACGATGGAGG 47413
QY 821 TGCCCGCAGAGGACGAGGTACAATCTTATCCTTCCAAGCAGCACTATGTGTAATGCTC 880
Db 47412 TGCCCGCAGAGGACGAGGTACAATCTTATCCTTCCAAGCAGCACTATGTGTAATGCTC 47353
QY 881 TAAGACCTCTCAGCACGGGGAGAAACTCCCGGAGAGCTCACCCCAAAACAAGGAGA 940
Db 47352 TAAGACCTCTCAGCACGGGGAGAAACTCCCGGAGAGCTCACCCCAAAACAAGGAGA 47293
QY 941 TCCCATCTAGATTCTTCTTGCTTTGACTCAGCTGGAAGTTAGAAAGCCTCGATTT 1000
Db 47292 TCCCATCTAGATTCTTCTTGCTTTGACTCAGCTGGAAGTTAGAAAGCCTCGATTT 47233
QY 1001 CATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATA 1060

Db 47232 CATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATA 47173
QY 1061 AAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCTCTATTCT 1120
Db 47172 AAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCTCTATTCT 47113
QY 1121 TTTTAAATATAAATCTTCTACTCTGTAGAGAGAAATGTGGTTTTAAATCTCTCTCACAT 1180
Db 47112 TTTTAAATATAAATCTTCTACTCTGTAGAGAGAAATGTGGTTTTAAATCTCTCTCACAT 47053
QY 1181 TTTGATGATTTAGACAGACTCCCCCTCTTCTCTCTAGTCAATAAACCCTTGTATCTA 1240
Db 47052 TTTGATGATTTAGACAGACTCCCCCTCTTCTCTCTAGTCAATAAACCCTTGTATCTA 46993
QY 1241 TTTCCAGCTTATCCCAAGAAAACTTTTGAAGGAAAGAGTAGACCCAAAGATGTTATT 1300
Db 46992 TTTCCAGCTTATCCCAAGAAAACTTTTGAAGGAAAGAGTAGACCCAAAGATGTTATT 46933
QY 1301 TTCTGCTGTTTGAATTTGTCTCCCCACCCCAACTTTGGCTAGTAAATAAACACTTACTGA 1360
Db 46932 TTCTGCTGTTTGAATTTGTCTCCCCACCCCAACTTTGGCTAGTAAATAAACACTTACTGA 46873
QY 1361 AGAAGAAGCAATAAGAGAAAGATATTTGTAAATCTCTCCAGCCCATGATCTCGGTTTCTT 1420
Db 46872 AGAAGAAGCAATAAGAGAAAGATATTTGTAAATCTCTCCAGCCCATGATCTCGGTTTCTT 46813
QY 1421 ACACCTGTGATCTTAAAGTTACCAACCAAAAGTCATTTTCAGTTTGGGCAACCAACCT 1480
Db 46812 ACACCTGTGATCTTAAAGTTACCAACCAAAAGTCATTTTCAGTTTGGGCAACCAACCT 46753
QY 1481 TTCTACTGCTGTTGACATCTTCTTATTACAGCAACACCATTTCTAGGAGTTTCTGAGCTC 1540
Db 46752 TTCTACTGCTGTTGACATCTTCTTATTACAGCAACACCATTTCTAGGAGTTTCTGAGCTC 46693
QY 1541 TCCACTGGAGTCCCTCTTCTGTGCGGGTCAGAAATGTCCCTAGATGAATGAGAAAATT 1600
Db 46692 TCCACTGGAGTCCCTCTTCTGTGCGGGTCAGAAATGTCCCTAGATGAATGAGAAAATT 46633
QY 1601 ATTTTAAATTAAGTCTCTAAATATAGTTAAATAAATAATGTTTTAGTAAATGATA 1660
Db 46632 ATTTTAAATTAAGTCTCTAAATATAGTTAAATAAATAATGTTTTAGTAAATGATA 46573
QY 1661 CACTATCTCTGTGAATAGCTCACCCCTACATGTGGATAGAGAAATGAAAAATAAT 1720
Db 46572 CACTATCTCTGTGAATAGCTCACCCCTACATGTGGATAGAGAAATGAAAAATAAT 46513
QY 1721 TGCTTTGACATTTGCTATATGCTACTTTGTAAAGTCATGCTTAAGTACAAATCCATGAA 1780
Db 46512 TGCTTTGACATTTGCTATATGCTACTTTGTAAAGTCATGCTTAAGTACAAATCCATGAA 46453
QY 1781 AA----- 1782
Db 46452 AAGCTCACTGATCCTAAATCTTTCCCTTTGAGGTCTCTATGGCTCTGATTGTACATGATA 46393
QY 1783 -----GCTCACACCTGTAAT 1797
Db 46392 GTAAGTGAAGCCATGTAAAGTAAATAATGTCTGGGCACAGTGGCTCACGCCTGTAAT 46333
QY 1798 CCTAGCACTTTGGGAGGCTGAGGAGGAGGATCACTTTGAGCCCAAGATTCGAGACTAGC 1857
Db 46332 CCTAGCACTTTGGGAGGCTGAGGAGGAGGATCACTTTGAGCCCAAGATTCGAGACTAGC 46273
QY 1858 CTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCA 1917
Db 46272 CTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCA 46213
QY 1918 TGGTGGCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAG 1977
Db 46212 TGGTGGCCTACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAG 46153
QY 1978 CCCAGGAGGTTGGGCTGCAGTGAAGCCATGATCACACCTGCACCTCCAGCCAGGTGAC 2037
Db 46152 CCCAGGAGGTTGGGCTGCAGTGAAGCCATGATCACACCTGCACCTCCAGCCAGGTGAC 46093


```
misc_feature 46985..55777
              /note="assembly_fragment"
misc_feature 55878..63178
              /note="assembly_fragment"
misc_feature 63279..73509
              /note="assembly_fragment"
misc_feature 73610..82711
              /note="assembly_fragment"
misc_feature 82812..93642
              /note="assembly_fragment"
misc_feature 93743..109313
              /note="assembly_fragment"
misc_feature 109414..124437
              /note="assembly_fragment"
misc_feature 124538..145477
              /note="assembly_fragment"

ORIGIN
Query Match      51.1%; Score 1083.2; DB 2; Length 145477;
Best Local Similarity 99.7%; Pred. No. 1.4e-236;
Matches 1085; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 701 CAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTGGCTTACAGCCTGGAGG 760
Db 18143 CAGCTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTGGCTTACAGCCTGGAGG 18202

QY 761 CTTCAAGCCAGCACTGGCTTTGGTCCAAACACCAAAACAAAGAGATATACGATGGAGG 820
Db 18203 CTTCAAGCCAGCACTGGCTTTGGTCCAAACACCAAAACAAAGAGATATACGATGGAGG 18262

QY 821 TGCCGCGACAGAGGACGAGGTACAATCTTATCCTTCCAAGCAGCACTATGTGTAATGCTC 880
Db 18263 TGCCGCGACAGAGGACGAGGTACAATCTTATCCTTCCAAGCAGCACTATGTGTAATGCTC 18322

QY 881 TAAGACCTCTCAGCACGGGCGGAAGAAACTCCCGAGAGCTCACCCAAACAAAGGAGA 940
Db 18323 TAAGACCTCTCAGCACGGGCGGAAGAAACTCCCGAGAGCTCACCCAAACAAAGGAGA 18382

QY 941 TCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTT 1000
Db 18383 TCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTT 18442

QY 1001 CATCTTTGAGAGGCCAAATGGTCTTAGCCTCAGCTCTCTCTCTCTCTCTCTCTACAT 1060
Db 18443 CATCTTTGAGAGGCCAAATGGTCTTAGCCTCAGCTCTCTCTCTCTCTCTCTCTACAT 18502

QY 1061 AAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTCT 1120
Db 18503 AAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTCT 18562

QY 1121 TTTTAAATATAAATTTCTACTCTGATGAGAGAAATGGTGGTTTAACTCTCTCTCTCAT 1180
Db 18563 TTTTAAATATAAATTTCTACTCTGATGAGAGAAATGGTGGTTTAACTCTCTCTCTCAT 18622

QY 1181 TTTGATGATTTAGACAGACTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTA 1240
Db 18623 TTTGATGATTTAGACAGACTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTA 18682

QY 1241 TTTCCCGAGCTATCCCCAAGAAACTTTTGAAGGAAAGAGTAGACCCCAAGATGTTATT 1300
Db 18683 TTTCCCGAGCTATCCCCAAGAAACTTTTGAAGGAAAGAGTAGACCCCAAGATGTTATT 18742

QY 1301 TTCTGCTGTTGAATTTTGTCTCCCAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTGA 1360
Db 18743 TTCTGCTGTTGAATTTTGTCTCCCAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTGA 18802

QY 1361 AGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTCTT 1420
Db 18803 AGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTCTT 18862

QY 1421 ACACGTGTGATCTTTAAAGTTACCAACCAAGTCAATTTTCAGTTTGAGGCAACCAACT 1480
Db 18863 ACACGTGTGATCTTTAAAGTTACCAACCAAGTCAATTTTCAGTTTGAGGCAACCAACT 18922
```

```
QY 1481 TTCTACTGCTGTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTC 1540
Db 18923 TTCTACTGCTGTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTC 18982

QY 1541 TCCACTGGAGTCCTTCTTCTGCGCGGTGAGAAATTTGCTTCCCTAGATGAATGAGAAAATT 1600
Db 18983 TCCACTGGAGTCCTTCTTCTGCGCGGTGAGAAATTTGCTTCCCTAGATGAATGAGAAAATT 19042

QY 1601 ATTTTAAATTTAAGTCCTAAATATAGTTAAATAAATAATGTTTAAATAAATAATGATA 1660
Db 19043 ATTTTAAATTTAAGTCCTAAATATAGTTAAATAAATAATGTTTAAATAAATAATGATA 19102

QY 1661 CACTATCTCTGTGAATAGCCTCACCCCTACATGTGGATAGAAGGAATGAAAAATAAT 1720
Db 19103 CACTGTCTCTGTGAATAGCCTCCCCCTACATGTGGATAGAAGGAATGAAAAATAAT 19162

QY 1721 TGCTTTGACATTTGCTATATATGGTACTTTTAAAGTCATGCTTAAGTACAAATTCATGAA 1780
Db 19163 TGCTTTGACATTTGCTATATATGGTACTTTTAAAGTCATGCTTAAGTACAAATTCATGAA 19222

QY 1781 AAGCTCAC 1788
Db 19223 AAGCTCAC 19230
```

```
RESULT 11
BD237992
LOCUS BD237992 982 bp DNA linear PAT 17-JUL-2003
DEFINITION Gastric polypeptide ZSIG28.
ACCESSION BD237992
VERSION BD237992.1 GI:33047762
KEYWORDS JP 2002524103-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 982)
AUTHORS Sheppard,P.O. and Foley,K.P.
TITLE Gastric polypeptide ZSIG28
JOURNAL Patent: JP 2002524103-A 1 06-AUG-2002;
ZYMOGENETICS INC
```

```
COMMENT
OS Homo sapiens (human)
PN JP 2002524103-A/1
PD 06-AUG-2002
PF 14-SEP-1999 JP 2000570197
PR 16-SEP-1998 US 09/154444
PI PAUL O SHEPPARD,KEVIN P FOLEY
PC C12N15/09,A61K38/00,A61K39/395,A61P1/04,A61P1/14,
PC A61P3/08,
PC A61P5/50,A61P31/04,A61P31/10,A61P35/00,C07K14/47,C07K16/18, PC
C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12P21/08,C12Q1/02 PC
,C12Q1/68,G01N33/15,
PC G01N33/50,G01N33/53,G01N33/577,C12N15/00,C12N5/00,A61K37/02 CC
Gastric polypeptide ZSIG28
FH Key Location/Qualifiers
FT CDS (70)..(853).
```

```
FEATURES
source
1..982
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

```
ORIGIN
Query Match      45.0%; Score 954.2; DB 6; Length 982;
Best Local Similarity 99.1%; Pred. No. 1.6e-207;
Matches 970; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 21 TTAGCTTCACACCTTCGGCAGCAGGAGGGCGGAGCTTCTCGAGGGCGGCGGGCG 80
Db 1 TTAGCTTCACTCTTCGGCAGCAGGAGGGCGGAGCTTCTCGAGGGCGGCGGGCG 60
```



```
QY      81  GCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGGTTCTCTCTGTCCATCCTGGGG 140
Db      61  GCCAGTATCATGTCCACCACCACATGCCAAGTGGTGGGTTCTCTCTGTCCATCCTGGGG 120

QY     141  CTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGAC 200
Db     121  CTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGAC 180

QY     201  AACCCCGTCACTCCGTGTTCCAGTACGAAGGCTCTGGAGAGCTGCGTGAGGCAGAGT 260
Db     181  AACCCCGTCACTCCGTGTTCCAGTACGAAGGCTCTGGAGAGCTGCGTGAGGCAGAGT 240

QY     261  TCAGGCTTCAACCGAATGCAGGCCCTATTTCACCATCCTGGGACTTCCAGCCATGCTGCAG 320
Db     241  TCAGGCTTCAACCGAATGCAGGCCCTATTTCACCATCCTGGGACTTCCAGCCATGCTGCAG 300

QY     321  GCAGTGGAGCCCTGATGATCGTAGGCATCGTCTCTGGGTGCCATTGGCCTCCTGGTATCC 380
Db     301  GCAGTGGAGCCCTGATGATCGTAGGCATCGTCTCTGGGTGCCATTGGCCTCCTGGTATCC 360

QY     381  ATCTTTGCCCTGAAATGCATCCGATTCGGAGTGGCAGCATGGAGGACTCTGCCAAAGCCAAATG 440
Db     361  ATCTTTGCCCTGAAATGCATCCGATTCGGAGTGGCAGCATGGAGGACTCTGCCAAAGCCAAATG 420

QY     441  ACACTGACCTCCGGGATCATGTTCAATTGTCTCAGGTCTTTGTGCAATTGCTGAGTGTCT 500
Db     421  ACACTGACCTCCGGGATCATGTTCAATTGTCTCAGGTCTTTGTGCAATTGCTGAGTGTCT 480

QY     501  GTGTTTGCCAAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCCGC 560
Db     481  GTGTTTGCCAAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCCGC 540

QY     561  ATGGGTGGGATGGTGACACTGTTTCAGACCAGGTACACATTTGGTGGGCTCTGTTTCGTG 620
Db     541  ATGGGTGGGATGGTGACACTGTTTCAGACCAGGTACACATTTGGTGGGCTCTGTTTCGTG 600

QY     621  GGTGTTGCTCGCTGGAGCCCTCACACTAATTGGGGGTGTGATGATGTGCAATCGCCTGCCGG 680
Db     601  GGTGTTGCTCGCTGGAGCCCTCACACTAATTGGGGGTGTGATGATGTGCAATCGCCTGCCGG 660

QY     681  GGCCTGGCACCAAGAACCAACTAATAAGCCGTTTCTATCATGCTCAGGCCACAGT 740
Db     661  GGCCTGGCACCAAGAACCAACTAATAAGCCGTTTCTATCATGCTCAGGCCACAGT 720

QY     741  GTTGCTTACAAGCCTGAGGCTTCAAGGCCAGCACTGGCTTTGGTCCAAACACCAAAAC 800
Db     721  GTTGCTTACAAGCCTGAGGCTTCAAGGCCAGCACTGGCTTTGGTCCAAACACCAAAAC 780

QY     801  AAGAAGATATACGATGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAG 860
Db     781  AAGAAGATATACGATGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAG 840

QY     861  CACGACTATGTGTAATGCTCTAAGACCTCTCAGCACGGGGCGGAAGAAACTCCCGGAGAGC 920
Db     841  CACGACTATGTGTAATGCTCTAAGACCTCTCAGCACGGGGCGGAAGAAACTCCCGGAGAAC 900

QY     921  TCACCCAAAAAACAGGAGATCCCATCTAGATTCTTCTTGTGCTTTGACTCACAGCTGGA 980
Db     901  TCACCCAAAAAACAGGAGATCCCATCTAGA-TTCTTCTTGTGCTTTGACTCACAGCTGGA 959

QY     981  AGTTAGAAAAGCCTCGATT 999
Db     960  AGTTAGAAAACCTCGAATT 978
```

```
RESULT 12
HSM800576/c
LOCUS      HSM800576      1171 bp      mRNA      linear      PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA; cDNA DKFZp564B2062 (from clone DKFZp564B2062).
ACCESSION  AL080075
VERSION    AL080075.1  GI:5262485
KEYWORDS   Homo sapiens (human)
```

```
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1171)
AUTHORS  Blum,H., Bauersachs,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE    Direct Submision
JOURNAL  Submitted (16-JUN-1999) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
COMMENT  Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by LMU (Ludwig Maximilians University,
Munich/Germany) within the cDNA sequencing consortium of the German
Genome Project.
This clone (DKFZp564B2062) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source           1..1171
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp564B2062"
                     /tissue_type="brain"
                     /clone_lib="564 (synonym: hfbr2). Vector pAMP1; host
                     Xl-2blue; sites NotI + SalI"
     polyA_signal     1130..1135
     polyA_site       1150
ORIGIN
```

```
Query Match      39.9%;      Score 845.4;  DB 9;      Length 1171;
Best Local Similarity 99.3%;      Pred. No. 1.3e-182;
Matches 849; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      701  CAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTGCTTACAAAGCCTGGAGG 760
Db      885  CAGCTACAAAGCCGTTTCTTATCATGCCTCGGCCACAGTGTGCTTACAAAGCCTGGAGG 826

QY      761  CTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAACAAGAGATATACGATGGAGG 820
Db      825  CTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAACAAGAGATATACGATGGAGG 766

QY      821  TGCCCCGCACAGAGGACGAGGTACAACTTTATCTCTTCAAGCACGACTATGTGTAATGCTC 880
Db      765  TGCCCCGCACAGAGGACGAGGTACAACTTTATCTCTTCAAGCACGACTATGTGTAATGCTC 706

QY      881  TAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCACCCAAAAACAAGGAGA 940
Db      705  TAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCACCCAAAAACAAGGAGA 646

QY      941  TCCCATCTAGATTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTT 1000
Db      645  TCCCATCCAGATTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTT 586

QY     1001  CATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCACCATTA 1060
Db     585  CATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCACCATTA 526

QY     1061  AACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTCT 1120
Db     525  AACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTCT 466

QY     1121  TTTTAAATATAAATCTTCTACTCTGATGAGAGAATGTGGTTTAAATCTCTCTCACAT 1180
Db     465  TTTTAAATATAAATCTTCTACTCTGATGAGAGAATGTGGTTTAAATCTCTCTCACAT 406

QY     1181  TTTGATGATTTAGACAGACTCCCGCTCTTCTCCTAGTCAATAAACCCATTGATGATCTA 1240
Db     405  TTTGATGATTTAGACAGACTCCCGCTCTTCTCCTAGTCAATAAACCCATTGATGATCTA 346

QY     1241  TTTCCCGCTTATCCCCAAGAAAACCTTTTGAAAGGAAGAGTAGACCCAAAGATGTTATT 1300
```

| | | | | |
|----|------|--|--------------------------------------|-----|
| Db | 345 | TTTCCCAGCTTATCCCCAAGAAAAC | TTTTTGAAAGGAAAGAGTAGACCCCAAGATGTTATT | 286 |
| QY | 1301 | TTCTGCTGTTTGAAATTTTGCTCCCCACCCCAACTTGGCTAGTAATAACACTTACTGA | 1360 | |
| Db | 285 | TTCTGCTGTTTGAAATTTTGCTCCCCACCCCAACTTGGCTAGTAATAACACTTACTGA | 226 | |
| QY | 1361 | AGAAGAAGCAATAAGAGAAAAGATATTTGTAATCTCTCCAGCCCCATGATCTCGGTTTCTT | 1420 | |
| Db | 225 | AGAAGAAGCAATAAGAGAAAAGATATTTGTAATCTCTCCAGCCCCATGATCTCGGTTTCTT | 166 | |
| QY | 1421 | ACACTGTGATCTTAAAGTTACCAAAACCAAGTCATTTTCAGTTTGAGGCAACCAAACTT | 1480 | |
| Db | 165 | ACACTGTGATCTTAAAGTTACCAAAACCAAGTCATTTTCAGTTTGAGGCAACCAAACTT | 106 | |
| QY | 1481 | TTCTACTGCTGTTGACATCTTCTTATTACAGCAACACCACTTCTAGGAGTTTCTCTGAGCTC | 1540 | |
| Db | 105 | TTCTACTGCTGTTGACATCTTCTTATTACAGCAACACCACTTCTAGGAGTTTCTCTGAGCTC | 46 | |
| QY | 1541 | TCCACTGGAGTCCTC | 1555 | |
| Db | 45 | TCCACTGGAGTCCTC | 31 | |

RESULT 13
AF221068
LOCUS AF221068 2786 bp mRNA linear ROD 10-OCT-2001
DEFINITION Mus musculus Claudin-18 mRNA, complete cds.

| ORIGIN | | Query Match | 39.5%; | Score 837.2; | DB 10; | Length 2786; |
|--------|----|--|--------------|---------------------|------------|--------------|
| | | Best Local Similarity | 73.4%; | Pred. No. 1.2e-180; | | |
| | | Matches 1300; | Conservative | 0; Mismatches 418; | Indels 52; | Gaps 16; |
| QY | 36 | CGGCAGCAGGAGGGCGGCAGCTTCTTCGCAGGCGGCAGGGCGGCCAGGATCATGTCC | 95 | | | |
| | | | | | | |
| ph | 16 | CAGCAGACAGCGGTGGTGCTCTCTGGAGGCTGC--GCCCAGCAGCTGAAGCCATGGCC | 73 | | | |

| | | | |
|----|------|---|------|
| QY | 96 | ACCACCAATGCCAAGTGGTGGGTTCTCTCTGTCCATCCTGGGGCTGGCCGGCTGCATC | 155 |
| Db | 74 | ACCAACAGTGCAGGTGGTAGGGCTTCTCCTGTCCCTCTGGGTCTGGCCGGCTGCATA | 133 |
| QY | 156 | GCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACCTCC | 215 |
| Db | 134 | GCCGCCACTGGGATGGACATGTGGAGCACTCAAGACCTGTATGACAACCCAGTCAACGCC | 193 |
| QY | 216 | GTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTAGGCGAGAGTTCAGGCTTCAACGAA | 275 |
| Db | 194 | GTGTTCCAGTATGAAGGGCTCTGGAGGAGTTGCGTGCAACAGAGCTCGGGGTTACCGAG | 253 |
| QY | 276 | TGCAGGGCCCTATTTCAACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTG | 335 |
| Db | 254 | TGCCGGCCATATTCAACCATCCTGGGCTTCCAGCCATGCTGCAAGCTGTACGAGCCCTG | 313 |
| QY | 336 | ATGATCGTAGGCATCGTCTGGGTGCCATTTGGCCTCCTGGTATCCATCTTTGCCCTGAAA | 395 |
| Db | 314 | ATGATCGTGGGCATGTTCTGGGGTTCATCGGTATCCTCGTGTCCATCTTCGCCCTGAAG | 373 |
| QY | 396 | TGCATCCGCATTTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACACTGACCTCCGG | 455 |
| Db | 374 | TGCATTCGCATTTGGTAGCATGGATGACTCTGCCAAGGCCAAGATGACTCTGACTTCTGG | 433 |
| QY | 456 | ATCATGTTCTATGTTCTCAGGTCTTTGTGCAATGCTGGAGTGTGTGTTGCCAACATG | 515 |
| Db | 434 | ATCTTGTTTCATCATCTCCGGCATCTGTGCAATCATTTGGTGTGTGTGTGTTGCCAACATG | 493 |
| QY | 516 | CTGGTGACTAACTTCTGGATGTCCACAGCTAAACATGTACA-----CCGGCATGGGT | 566 |
| Db | 494 | CTGGTGACCAACTTCTGGATGTCCACAGCTAAACATGTACAGCGGCATGGCGGCATGGGT | 553 |
| QY | 567 | GGGATGGTGAGACTGTTTCAGACCAGGTACACATTTGGTGGGCTCTGTTCGTGGGCTGG | 626 |
| Db | 554 | GGCATGGTGACAGACCGTTCAGACCAGGTACACCTTCGGTGCAGCTCTGTTCGTGGGCTGG | 613 |
| QY | 627 | GTCGCTGGAGGCTCAACATAATTGGGGGTGTGATGTGCATCGCCTGCCGGGGCCCTG | 686 |
| Db | 614 | GTTGCTGGAGGCCCTACCCCTGATTGGGGAGTGTGATGTGCATCGCCTGCGGTGGCCTG | 673 |
| QY | 687 | GCACAGAGAAGAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTGTC | 746 |
| Db | 674 | ACACAGATGACAGCAACTTCAAAGCTGTGTCTTACCATGCCTCTGGCCAAAATGTTGCC | 733 |
| QY | 747 | TACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAAACAAGAAG | 806 |
| Db | 734 | TACAGGCTGGAGGCTTTAAGGCCAGCACTGGCTTTGGGTCCAACACCAAGAAACAAGAAG | 793 |
| QY | 807 | ATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGAC | 866 |
| Db | 794 | ATCTACGATGGGGTGCCTGCACAGAACGATGAACAGTCTCATCTCTACCAAGTATGAC | 853 |
| QY | 867 | TATGTGTAATGCTCTAAGACTCTCAGCACGGGGCGAAGAACTCCC-----GGAGA | 918 |
| Db | 854 | TATGTGTAGTGCTCTAAGACCGGCCAACCTGTGTGCAGGAGGAACCCCTTCCCAAGAAGA | 913 |
| QY | 919 | GCTCACCAAAAAACAAGGAGATCCCATCTAGATTTCTTCTTGTCTTTGACTCACAGCTG | 978 |
| Db | 914 | GCTCACCCCAAAG-CAACGGGAGTCTACCTTGTTCCTTGTGATTTCAACTGACATCTG | 972 |
| QY | 979 | GAAGTTAGAAAAAGCCTCGATTTTCATCTTTGGAGAGGCCAAAATGGTCTTAGCCTCA-GTCT | 1037 |
| Db | 973 | AAAGTTGGTAAGCCTGATTTTCATCCATAGGGAGGTAGACAGTCTTGGCCACATGTGT | 1032 |
| QY | 1038 | CTGTCTCTAAATATTCCACCATAAAACAGCTGAGTTA--TTTATGAATTAGAGGCTATA | 1094 |
| Db | 1033 | CTGCCCTCTAAATATCCCATCACAAAAACAGCTGAGTTATCGTTTATGAGTTAGAGGCCATA | 1092 |
| QY | 1095 | GCTCAATTTTCAATCCTCTATTTCTTTTTTTTAAATATAACT--TTCTACTCTGATGAGA | 1152 |
| Db | 1093 | ACACTCACTTAGGCCAAACCTCTGCTTTTTTACCGTAGACTTTCTTTTTCATCTGGTGATG | 1152 |

```
QY 1153 GAATGTGGTTTAAATCTCTCTCACATTTTGATGATTTAGACAGACTCCCCTCTTCT 1212
Db 1153 GAATGGAATTTGACTCACAGACTAATACTTTAATGGTTTATAGAGAACTTTC---CTTCT 1209

QY 1213 CCTAGTCAATAAACCCATTGATGATCTAATTTCCAGCTT-ATCCCCAAGAAAACTTTGA 1271
Db 1210 CGTACTTAATAAGCCTGCTGATGGTCGATTTCCAGCTTGACCACCAAGGAAATTTAA 1269

QY 1272 AAGGAAAGAGTAGACCCCAAGATGTTATTTCTGCTGTTTGAATTTGTCTCCCCACCCC 1331
Db 1270 AAGGAAAAAATAACATTAAGGCAATATTCTCTACTCAATGTGCTTACCACCCC 1329

QY 1332 CAACTTGGCTAGTAATA-----AACACTTACTGAAGAGCAATAAGAGAAATAT 1385
Db 1330 CAACTTGACTGATAATAATAATGAACACCACCTTAAGAGAAAGATGCCAGAGGAAGATAG 1389

QY 1386 TTGT-AACTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATCTTAAAGTTACCA 1444
Db 1390 TTGTGTTTCCCCCAGCCAGTCATCTGAGTCCCCCTATGTGGTGATCTAGAACATTA-CT 1448

QY 1445 AACCAAGTCATTTTCAGTTTGGGCAACCAACCTTTTACTGCTGTGTGACATCTTCTT 1504
Db 1449 CGCCACAGTGATTTTCA-AAGAGGCAAGCGAGCCTGTTCGCTCTGCTCAGCATCTGCTG 1507

QY 1505 ATTACAGCAACACCATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTCTGTCG 1564
Db 1508 ATCCAGCAAGGCCCTTCCAGAGCTTTCCA-----CTAGAAGTCTCTCTCTCTCG 1558

QY 1565 CGGTCAGAAATTTGCCCTAGATGAATGAGAAA---ATTATTTTTTTAAATTAAGTCCT 1621
Db 1559 GAAGTCAGAAATTTCCCCCTAGAGAGTAAGAAATAGATTCTTTTGGGTAACTGAGTCCT 1618

QY 1622 AAATATAGTTAAATAAATAATGTTTGTAGTAAATGATACACTATCTCTGTGAAATAGCC 1681
Db 1619 AGGTATAGTTATAATAAATAGTATATTAGCAAAACGGTTTGGTATCTCAGTGAATTAGTT 1678

QY 1682 TCACCCCTACATGTGGATAGA-AGGAATGAAAAATAAATTGCTTTGACATGTCTATAT 1740
Db 1679 TCAGCCTTACATATAGAAAAAGCTGGGAAAAAAGCAATCCCTTGACATGTCTATAG 1738

QY 1741 GGTACTTTGTAAAGTCATGCTTAAGTACAA 1770
Db 1739 CGTAAGATCCTATATAAATCCAAGCTTCAA 1768
```

```
RESULT 14
AF221069
LOCUS AF221069 Homo sapiens Claudin-18 mRNA, complete cds. PRI 10-OCT-2001
DEFINITION AF221069 Homo sapiens Claudin-18 mRNA, complete cds.
ACCESSION AF221069.1 GI:6715517
VERSION AF221069.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 816) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B., Popescu,N.C. and Kimura,S.
TITLE claudin-18, a novel downstream target gene for the T/EBP/NKX2.1 homeodomain transcription factor, encodes lung- and stomach-specific isoforms through alternative splicing
JOURNAL Mol. Cell. Biol. 21 (21), 7380-7390 (2001)
MEDLINE 21470315
PUBMED 11585919
REFERENCE 2 (bases 1 to 816) Niimi,T. and Kimura,S.
AUTHORS Direct Submission
TITLE Submitted (04-JAN-2000) Laboratory of Metabolism, DBS, National Cancer Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
JOURNAL Location/Qualifiers
FEATURES
source 1. .816
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
10..795
/codon_start=1
/product="Claudin-18"
/protein_id="AAF26448.1"
/db_xref="GI:6715518"
/translation="MSTTTCQVAVFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVF
QYGLWRSQVRSQSGFTFCRPYFTILGLPAMLQAVRALMIVGLVGLLVSIFALK
CIRIGMEDSAKANMTLTSIMFIVSGLCAIAGVSFANMLVTNFMSTANMTYTMGG
MVQTVQTRYTFGAALFVGVAGGLTLIGVNMCIACRGLAPEETNYKAVSYHASGHSV
AYKPGGFKASTGFGSNTKNKKIYDGGARTEDVQSYPSKHDYV"
```

CDS

ORIGIN

```
Query Match 38.5%; Score 816; DB 9; Length 816;
Best Local Similarity 100.0%; Pred. No. 6.5e-176;
Matches 816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GCCAGGATCATGTCCACCACCATGCCAAGTGGTGGCGTTCCTCCTGTCCATCCTGGGG 140
Db 1 GCCAGGATCATGTCCACCACCATGCCAAGTGGTGGCGTTCCTCCTGTCCATCCTGGGG 60

QY 141 CTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGAC 200
Db 61 CTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGAC 120

QY 201 AACCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGCAGAGT 260
Db 121 AACCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGCAGAGT 180

QY 261 TCAGGCTTCAACCGAATGCAGGCCCTATTTCACCATCTCTGGGACTTCCAGCCATCTGCAG 320
Db 181 TCAGGCTTCAACCGAATGCAGGCCCTATTTCACCATCTCTGGGACTTCCAGCCATCTGCAG 240

QY 321 GCAGTGCAGGCCCTGATGATCGTAGGCATCGTCTCTGGGTGCCATTGGCCCTCTGGTATCC 380
Db 241 GCAGTGCAGGCCCTGATGATCGTAGGCATCGTCTCTGGGTGCCATTGGCCCTCTGGTATCC 300

QY 381 ATCTTTGCCCTGAAATGCATCCGCATTGGCAGATGGAGGACTCTGCCAAAGCCAACATG 440
Db 301 ATCTTTGCCCTGAAATGCATCCGCATTGGCAGATGGAGGACTCTGCCAAAGCCAACATG 360

QY 441 ACACCTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGTGGAGTGTCT 500
Db 361 ACACCTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGTGGAGTGTCT 420

QY 501 GTGTTTGCCAAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGC 560
Db 421 GTGTTTGCCAAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGC 480

QY 561 ATGGTGGGATGGTGACAGACTGTTTCAGACCAAGGTACACATTTGGTGGGCTCTGTTCTGTG 620
Db 481 ATGGTGGGATGGTGACAGACTGTTTCAGACCAAGGTACACATTTGGTGGGCTCTGTTCTGTG 540

QY 621 GGCTGGGTGCTGGAGGCCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGG 680
Db 541 GGCTGGGTGCTGGAGGCCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGG 600

QY 681 GGCCTGGCACCAGAAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGT 740
Db 601 GGCCTGGCACCAGAAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGT 660

QY 741 GTTGCCCTACAAGCCTGGAGGCTTCAAGGCCAGCAGCTGGCTTTGGGTCCAACACCAAAAAC 800
Db 661 GTTGCCCTACAAGCCTGGAGGCTTCAAGGCCAGCAGCTGGCTTTGGGTCCAACACCAAAAAC 720

QY 801 AAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAACTTATCCTTCCAAG 860
Db 721 AAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAACTTATCCTTCCAAG 780

QY 861 CACGACTATGTGTAATGCTCTAAGACCTCTCAGCAC 896
Db 781 CACGACTATGTGTAATGCTCTAAGACCTCTCAGCAC 816
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 14:08:53 ; Search time 118 Seconds
(without alignments)
9975.007 Million cell updates/sec

Title: US-10-063-731-117
Perfect score: 2121
Sequence: 1 gagctcccttcaggagcg.....ggttaaaactaatcttttaa 2121

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query % | | Length | DB | ID | Description |
|------------|-------|---------|-------|--------|---------------------|----|-------------------|
| | | Match | | | | | |
| 1 | 733 | 34.6 | 1353 | 4 | US-09-724-864-32 | | Sequence 32, Appl |
| 2 | 174.4 | 8.2 | 39982 | 4 | US-09-820-924-3 | | Sequence 3, Appli |
| C 3 | 170.8 | 8.1 | 9365 | 4 | US-09-608-285A-8 | | Sequence 8, Appli |
| C 4 | 170.8 | 8.1 | 9365 | 4 | US-09-350-836B-8 | | Sequence 8, Appli |
| C 5 | 170.8 | 8.1 | 9365 | 4 | US-09-370-265-8 | | Sequence 8, Appli |
| C 6 | 170.8 | 8.1 | 9365 | 4 | US-09-557-800C-8 | | Sequence 8, Appli |
| C 7 | 170.8 | 8.1 | 9365 | 4 | US-09-370-625A-8 | | Sequence 8, Appli |
| C 8 | 170.8 | 8.1 | 14747 | 4 | US-09-608-285A-42 | | Sequence 42, Appl |
| C 9 | 170.8 | 8.1 | 14747 | 4 | US-09-557-800C-42 | | Sequence 42, Appl |
| C 10 | 170.8 | 8.1 | 15977 | 4 | US-09-608-285A-59 | | Sequence 59, Appl |
| C 11 | 170.2 | 8.0 | 44848 | 4 | US-09-435-739-42 | | Sequence 42, Appl |
| 12 | 170 | 8.0 | 685 | 4 | US-09-227-357-100 | | Sequence 100, App |
| 13 | 169.8 | 8.0 | 13205 | 4 | US-09-835-811-3 | | Sequence 3, Appli |
| 14 | 169.8 | 8.0 | 98844 | 4 | US-09-791-211-10 | | Sequence 10, Appl |
| C 15 | 169.6 | 8.0 | 63000 | 4 | US-09-780-172-18 | | Sequence 18, Appl |
| C 16 | 169.6 | 8.0 | 75395 | 4 | US-09-844-890-3 | | Sequence 3, Appli |
| C 17 | 167.8 | 7.9 | 63588 | 4 | US-09-873-404-3 | | Sequence 3, Appli |
| 18 | 166.8 | 7.9 | 17000 | 4 | US-09-679-299A-18 | | Sequence 18, Appl |
| C 19 | 165.4 | 7.8 | 81001 | 4 | US-09-750-580-1 | | Sequence 1, Appli |
| 20 | 164.8 | 7.8 | 6038 | 3 | US-09-305-639-4 | | Sequence 4, Appli |
| 21 | 164.8 | 7.8 | 6038 | 4 | US-09-525-160B-2 | | Sequence 2, Appli |
| 22 | 164.8 | 7.8 | 7622 | 3 | US-09-305-639-1 | | Sequence 1, Appli |
| 23 | 164.8 | 7.8 | 7622 | 4 | US-09-525-160B-1 | | Sequence 1, Appli |
| 24 | 164.4 | 7.8 | 436 | 4 | US-09-621-976-19010 | | Sequence 19010, A |
| 25 | 164.2 | 7.7 | 1519 | 4 | US-09-716-129-45 | | Sequence 45, Appl |
| 26 | 163.4 | 7.7 | 45546 | 4 | US-09-146-053-6 | | Sequence 6, Appli |
| C 27 | 163.4 | 7.7 | 99500 | 4 | US-09-798-096-10 | | Sequence 10, Appl |

ALIGNMENTS

RESULT 1

US-09-724-864-32
; Sequence 32, Application US/09724864

; Patent No. 6380362
; GENERAL INFORMATION:

;
;
;

```

; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
;
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
;
; FILE REFERENCE: 11000.1050U1

```

```

; CURRENT APPLICATION NUMBER: US/0
; CURRENT FILING DATE: 2000-11-28

```

; PRIOR APPLICATION NUMBER: U.S.
; PRIOR FILING DATE: 1999-12-23

```

; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSEQ for Windows Version 4.0

```

```

; SEQ ID NO 32
; LENGTH: 1353
; TYPE: DNA

```

Query Match 34.6%; Score 733; DB 4; Length 1353;
Best Local Similarity 76.9%; Pred. No. 3.9e-185;
Matches 1006; Conservative 0; Mismatches 275; Indels 28; Gaps 8.

| | | | |
|----|-----|---|-----|
| Qy | 1 | GAGCTCCCTCAGGAGCGGTTAGCTTACACACTTCGGCAGCAGGAGGGCGGCAGCTTCT | 60 |
| Db | 46 | | |
| | | GAGCGTCCCAGCGAGGCCGAGGCCGTGAACCTTCCAGCAAGAGGGTGTTGCT | 105 |
| Qy | 61 | CGCAGCGGCAG-GGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGTTGGCG | 119 |
| Db | 106 | | |
| | | CCTGGAAGCCTGCGCCAGCAGCTGAAGCCATGGCCACCACCGTGCAGGTGGTAGGG | 165 |
| Qy | 120 | TTCCTCTGTCCAATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGG | 179 |
| Db | 166 | | |
| | | CTTCTCTGTCCCTCCTGGTCTGGCCGGCTGCATAGCCGCCACTGGGATGGACATGTGG | 225 |
| Qy | 180 | AGCACCCAGGACCTGTACGACAACCCCGTCACTCCGTGTTCCAGTACGAAGGGCTCTGG | 239 |
| Db | 226 | | |
| | | AGCACTCAAGACCTGTATGACAACCCAGTCAACCGCGTGTCCAGCATGAAGGGCTCTGG | 285 |
| Qy | 240 | AGGAGCTCGGTAGGCAGAGTTTCAGGCTTCACCGAATGCAGGCCCTATTACCATCCTG | 299 |
| Db | 286 | | |
| | | AGGAGTTGCGTGCAACAGAGCTCGGGGTTCACCGAGTGCCGGCCATACTTACCATCCTG | 345 |
| Qy | 300 | GGACTTCCAGCCATGCTGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGTCCCTGGGT | 359 |
| Db | 346 | | |
| | | GGCCTTCCAGCCATGCTGCAAGCTGTACAGCCCCGTGATCGTGGGCATGTTCTGGGG | 405 |

QY 360 GCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGATTTGGCAGCATGGAG 419
Db 406 GTCATCGGTATCCTCGTGTCCATCTTCGCCCTGAAAGTGCAATTCGCATGGTAGCATGGAT 465
QY 420 GACTCTGCCAAAGCAACATGACACTGACCTCCGGGATCATGTTCAATGTCTCAGGCTTT 479
Db 466 GACTCTGCCAAGCAACATGACACTGACCTCTGGGATCTTGTTCATCATCTCCGGCATC 525
QY 480 TGTGCAATTTGCTGGAGTGTCTGTGTTTGCACATGCTGGTGACCTAACTTCTGGATGTC 539
Db 526 TGTGCAATCATTTGGTGTGTCTGTGTTTGCACATGCTGGTGACCAACTTCTGGATGTC 585
QY 540 ACAGCTAACATGTACA-----CCGGCATGGGTGGATGGTGACACTGTTTCAGACC 590
Db 586 ACAGCTAACATGTACAGCGGCATGGCGGCATGGGTGGCATGGTGACACCGTTTCAGACC 645
QY 591 AGGTACACATTTGGTGGGCTCTGTTCGTGGGTGGGTGGTGACCTGAGGCTCACACTAATT 650
Db 646 AGGTACACCTTCCGTGGAGTCTGTTCGTGGGTGGGTGGTGACCTGAGGCTCACCTGATT 705
QY 651 GGGGGTGTGATGATGTGCATCGCTGCGGGCTGGCCACCAAGAGATATACGATGGAGGTTCAGACA 710
Db 706 GGGGGAGTGATGATGTGCATCGCTGCGGGCTGGCCACCAAGAGATATACGATGGAGGTTCAGACA 765
QY 711 GCCGTTTCTTATCATGCTCAGGCCACAGTGTGCTTACAGCCTGGAGGTTCAAGGCC 770
Db 766 GCTGTGCTTACCATGCTCTGGCCAAATGTTGCTTACAGGCTGGAGGCTTAAAGGCC 825
QY 771 AGCACTGGCTTTGGGTCCAAACCAACCAAGAGATATACGATGGAGGTTCAGACA 830
Db 826 AGCACTGGCTTTGGGTCCAAACCAACCAAGAGATATACGATGGAGGTTCAGACA 885
QY 831 GAGGACGAGGTACAACTTATCTTCCAGCACGACTATGTGTAATGCTCTAAGACCTCT 890
Db 886 GAAGACGATGAACAGTCTCATCTTACCAAGTATGACTATGTGTAGTGTCTTAAAGACCCGC 945
QY 891 CAGCAGCGGGGGAAGAACTCCC-----GGAGAGCTCACCCCAACCAAGAGATC 942
Db 946 CAACCTGTGTGAGGAGGAACCTTCCCAAGAGAGCTCACCC-CAAAGCAAGGGAGT 1004
QY 943 CCATCTAGATTCTCTTCTGCTTTGACTCACAGCTGGAAGTAGAAGAGCCTCGATTCA 1002
Db 1005 CTACCTTGTTCCTTGTGATTTCAACTGACATCTGAAAGTGGTAAAGCCTGATTTC 1064
QY 1003 TCTTTGGAGAGGCCAAATGGTCTTAGCCTCA-GTCTCTGTCTTAAATATCCACCATAA 1061
Db 1065 TCCATAGGAGGCTAGACAGTCTTGGCCACATGTGTCTGCTCTTAAATATCCCATCAAA 1124
QY 1062 AACAGCTGAGTTA---TTTATGATTTAGAGGCTATAGCTCACATTTTCAATCTCTATT 1118
Db 1125 AACAGCTGAGTTATCGTTTATGATTTAGAGGCCATAACACTTACCTTAGCCCAACCTCT 1184
QY 1119 CTTTCTTAAATATAACT--TTCTACTCTGATGAGAGATGTGGTTTAACTCTCTCTC 1176
Db 1185 GCTTTTACCGTAGACTTTCTTTTCAATCTGGTGATGGAATGGAATTTGACTCACAGACTA 1244
QY 1177 ACATTTTGTATGATTTAGACAGACTCCCTCTTCTCTCTAGTCAATAAACCCATTGATGA 1236
Db 1245 ATACTTTAATGGTTTAGAGAACTTTC---CTTCTCTGACTTAAAGCCTGCTGATGG 1301
QY 1237 TCTATTTCCAGCTTATCCCAAGAAACTTTTGAAGGAAAGAGTAGA 1285
Db 1302 TCGATTTTCCAGCTTGACCACCAAGGGAATTTTAAAAAATAAATAAATAAATAAATAA 1350

RESULT 2
US-09-820-924-3
; Sequence 3, Application US/09820924
; Patent No. 655351
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001213
; CURRENT APPLICATION NUMBER: US/09/820,924
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 39982
; TYPE: DNA
; ORGANISM: Human
; US-09-820-924-3

Query Match 8.2%; Score 174.4; DB 4; Length 39982;
Best Local Similarity 77.6%; Pred. No. 1.3e-35;
Matches 239; Conservative 0; Mismatches 61; Indels 8; Gaps 2;

QY 1784 CTCACACCTGTATCTCTAGCAGCTTTGGGAGGCTGAGGAGGAGGATCACTTGAGCCCGAGA 1843
Db 31724 CTCATGCCTGTATCTCTAGCAGCTTTGGGAGGCTGAGGAGGAGGATCACTTGAGCCCGAGA 31783
QY 1844 ACTTCGAGACTAGCTTGGGCAACATGGAGAGGC-CTGTCTCTACAAATACAGAGAGAA 1902
Db 31784 AGTTTGAGACCAGCTTGGGCAACATGGGCAACCTCTGTCTCTACAAAGGATAC-----A 31836
QY 1903 AAAATCAGCCAGTCATGTGTGGCATAACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 1962
Db 31837 AAAATAGCTGGGTGTGTGTAGTGACACCTGAGTCCAGCTACCTGGGAGGCTGAGGTG 31896
QY 1963 GGAGGATCACTTGAGCCCGAGGAGTGGGGCTGCAGTGAGCCATGATCAACACCATGCA 2022
Db 31897 GGAGGATCACTTGAGCCCGAGGAGTGGGGCTGCAGTGAGCCATGATCAACACCATGCA 31956
QY 2023 CTCAGCCAGGTGACATAGCAGATCCTGTCTTAAATAAATAAATAAATAAATAAATAAATAA 2082
Db 31957 CTCAGCCAGGTGACATAGCAGATCCTGTCTTAAATAAATAAATAAATAAATAAATAAATAA 32016
QY 2083 CAGCAAGT 2090
Db 32017 AAGACATT 32024

RESULT 3
US-09-608-285A-8/c
; Sequence 8, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205

```
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(288)
; NAME/KEY: exon
; LOCATION: (1281)..(1580)
; NAME/KEY: exon
; LOCATION: (1820)..(1855)
; NAME/KEY: exon
; LOCATION: (2467)..(2555)
; NAME/KEY: exon
; LOCATION: (2863)..(2942)
; NAME/KEY: exon
; LOCATION: (3889)..(3950)
; NAME/KEY: exon
; LOCATION: (4894)..(4995)
; NAME/KEY: exon
; LOCATION: (5847)..(5987)
; NAME/KEY: exon
; LOCATION: (6966)..(7138)
; NAME/KEY: exon
; LOCATION: (8556)..(9365)
; NAME/KEY: misc feature
; LOCATION: (3409)
; OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
; NAME/KEY: misc feature
; LOCATION: (9214)
; OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
; NAME/KEY: misc feature
; LOCATION: (9303)
; OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
; NAME/KEY: misc feature
; LOCATION: (9311)
; OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
; US-09-608-285A-8

Query Match      8.1%; Score 170.8; DB 4; Length 9365;
Best Local Similarity 72.4%; Pred. No. 5.2e-35;
Matches 223; Conservative 26; Mismatches 53; Indels 6; Gaps 3;

QY 1783 GCTCACACCTGTAACTCTAGCACCTTTGGAGGCTGAGGAGGAGGATCCTTGAGCCCGAG 1842
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7714 GCTCAYRCCTRTATYCYAGCACCTTTGGAGGCTGAGGAGGAGGATCCTTGAGRTCAG 7655

QY 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAGCCCTGTCTCTACAAAATACAGAGAGAA 1902
Db      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7654 RAGTTYRAGAMYAGCCTGGCCAAACATGGTGAACCCCGTCTCTACTAAAAATACA---AA 7598

QY 1903 AAAATCAGCCAGTCATGGTGGCATAACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTG 1962
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7597 AAAATTAGCCRCGGCTGTGGCGCRYGCTGTAAATCCAGCTACTYGGGAGGCTGAGGCA 7538

QY 1963 GGAGGATCATTGAGCCCGAGGAGGTTGGGCTGCAGTGAGCCATGATCACACCCTGCA 2022
Db      |:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7537 GRAGAAATCGCTTGAAACCCA-GGAGGYGRAGGYTGCAATGAGCYGWGATCRYGCCAYTGCA 7479

QY 2023 CTCCAGCCAGGTGA--CATAGCGAGATCCTGTCTAAAAATAAAAAATAAATAATGAA 2080
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7478 CTCCAGCCYRGGMAACAAGACRAGACTCCCGTCTCAAAAAATAAATAAATAAATAAA 7419

QY 2081 CACAGCAA 2088
Db      |||||
7418 TAAATAAA 7411

RESULT 4
US-09-350-836B-8/c
; Sequence 8, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(288)
; NAME/KEY: exon
; LOCATION: (1281)..(1580)
; NAME/KEY: exon
; LOCATION: (1820)..(1855)
; NAME/KEY: exon
; LOCATION: (2467)..(2555)
; NAME/KEY: exon
; LOCATION: (2863)..(2942)
; NAME/KEY: exon
; LOCATION: (3889)..(3950)
; NAME/KEY: exon
; LOCATION: (4894)..(4995)
; NAME/KEY: exon
; LOCATION: (5847)..(5987)
; NAME/KEY: exon
; LOCATION: (6966)..(7138)
; NAME/KEY: exon
; LOCATION: (8556)..(9365)
; NAME/KEY: misc feature
; LOCATION: (3409)
; OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
; NAME/KEY: misc feature
; LOCATION: (9214)
; OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
; NAME/KEY: misc feature
; LOCATION: (9303)
; OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
; NAME/KEY: misc feature
; LOCATION: (9311)
; OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
; US-09-350-836B-8

Query Match      8.1%; Score 170.8; DB 4; Length 9365;
Best Local Similarity 72.4%; Pred. No. 5.2e-35;
Matches 223; Conservative 26; Mismatches 53; Indels 6; Gaps 3;

QY 1783 GCTCACACCTGTAACTCTAGCACCTTTGGAGGCTGAGGAGGAGGATCCTTGAGCCCGAG 1842
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7714 GCTCAYRCCTRTATYCYAGCACCTTTGGAGGCTGAGGAGGAGGATCCTTGAGRTCAG 7655

QY 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAGCCCTGTCTCTACAAAATACAGAGAGAA 1902
Db      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7654 RAGTTYRAGAMYAGCCTGGCCAAACATGGTGAACCCCGTCTCTACTAAAAATACA---AA 7598

QY 1903 AAAATCAGCCAGTCATGGTGGCATAACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTG 1962
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7597 AAAATTAGCCRCGGCTGTGGCGCRYGCTGTAAATCCAGCTACTYGGGAGGCTGAGGCA 7538

QY 1963 GGAGGATCATTGAGCCCGAGGAGGTTGGGCTGCAGTGAGCCATGATCACACCCTGCA 2022
Db      |:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7537 GRAGAAATCGCTTGAAACCCA-GGAGGYGRAGGYTGCAATGAGCYGWGATCRYGCCAYTGCA 7479

QY 2023 CTCCAGCCAGGTGA--CATAGCGAGATCCTGTCTAAAAATAAAAAATAAATAATGAA 2080
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7478 CTCCAGCCYRGGMAACAAGACRAGACTCCCGTCTCAAAAAATAAATAAATAAATAAA 7419

QY 2081 CACAGCAA 2088
Db      |||||
7418 TAAATAAA 7411
```

RESULT 5
US-09-370-265-8/c
; Sequence 8, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:


```
; NAME/KEY: misc feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-557-800C-8

Query Match      8.1%; Score 170.8; DB 4; Length 9365;
Best Local Similarity 72.4%; Pred. No. 5.2e-35;
Matches 223; Conservative 26; Mismatches 53; Indels 6; Gaps 3;

QY 1783 GCTCACACCTGTAATCCTAGCATTGGGAGGCTGAGGAGGAAGGATCACCTTGAGCCCGAG 1842
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
Db 7714 GCTCAYRCCTRTAATYCYAGCACTTTGGGAGGCGYAGGYGGGTGGATCACCTGAGRTGAG 7655

QY 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTCTCTACAAAATACAGAGAGAA 1902
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
Db 7654 RAGTTYRAGAMYAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACA---AA 7598

QY 1903 AAAATCAGCCAGTCATGCTGGGCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTG 1962
      ||||| |||||:| |||||:| |||||:| |||||:| |||||:| |||||:| |||||:|
Db 7597 AAAATAGCCRCGGCGTGGTGGCGCRYGCTGTATCCCAGCTACTYGGGAGGCTGAGGCA 7538

QY 1963 GGAGGATCACTTGAGCCCGAGGAGGTTGGGCTGAGTCCAGCATGATCACACCACTGCA 2022
      |:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7537 GRAGAATCGCTTGAACCCA-GGAGGYGRAGGYTGAGTGAGCYGWGATCRYGCCAYTGCA 7479

QY 2023 CTCCAGCCAGGTGA--CATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATGAA 2080
      ||||| |||||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7478 CTCCAGCCYRGGMAACAAGAGCRAGACTCCGCTCTCAAAAAAAATAATAATAAATAAA 7419

QY 2081 CACAGCAA 2088
      |||||
Db 7418 TAAATAAA 7411
```

```
RESULT 7
US-09-370-625A-8/c
; Sequence 8, Application US/09370625A
; Patent No. 6600032
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/09/370,625A
; CURRENT FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(288)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1281)..(1580)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1820)..(1855)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (2467)..(2555)
```

```
; FEATURE:
; NAME/KEY: exon
; LOCATION: (2863)..(2942)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (3889)..(3950)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (4894)..(4995)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (5847)..(5987)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (6966)..(7138)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (8556)..(9365)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3409)
; OTHER INFORMATION: n = a or g or t or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9214)
; OTHER INFORMATION: n = a or g or t or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9303)
; OTHER INFORMATION: n = a or g or t or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9311)
; OTHER INFORMATION: n = a or g or t or c
US-09-370-625A-8

Query Match      8.1%; Score 170.8; DB 4; Length 9365;
Best Local Similarity 72.4%; Pred. No. 5.2e-35;
Matches 223; Conservative 26; Mismatches 53; Indels 6; Gaps 3;

QY 1783 GCTCACACCTGTAATCCTAGCATTGGGAGGCTGAGGAGGAAGGATCACCTTGAGCCCGAG 1842
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
Db 7714 GCTCAYRCCTRTAATYCYAGCACTTTGGGAGGCGYAGGYGGGTGGATCACCTGAGRTGAG 7655

QY 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAA 1902
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
Db 7654 RAGTTYRAGAMYAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACA---AA 7598

QY 1903 AAAATCAGCCAGTCATGCTGGGCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTG 1962
      ||||| |||||:| |||||:| |||||:| |||||:| |||||:| |||||:| |||||:|
Db 7597 AAAATAGCCRCGGCGTGGTGGCGCRYGCTGTATCCCAGCTACTYGGGAGGCTGAGGCA 7538

QY 1963 GGAGGATCACTTGAGCCCGAGGAGGTTGGGCTGAGTCCAGCATGATCACACCACTGCA 2022
      |:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7537 GRAGAATCGCTTGAACCCA-GGAGGYGRAGGYTGAGTGAGCYGWGATCRYGCCAYTGCA 7479

QY 2023 CTCCAGCCAGGTGA--CATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATGAA 2080
      ||||| |||||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7478 CTCCAGCCYRGGMAACAAGAGCRAGACTCCGCTCTCAAAAAAAATAATAATAAATAAA 7419

QY 2081 CACAGCAA 2088
      |||||
Db 7418 TAAATAAA 7411
```

```
RESULT 8
US-09-608-285A-42/c
; Sequence 42, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
```

APPLICANT: Yeung, George

TITLE OF INVENTION: Methods and Materials Relating to CD39-Like

FILE REFERENCE: 28110/36457

CURRENT APPLICATION NUMBER: US/09/557,800C

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/481,238

PRIOR FILING DATE: 2000-01-11

PRIOR APPLICATION NUMBER: 09/370,265

PRIOR FILING DATE: 1999-08-09

PRIOR APPLICATION NUMBER: PCT/US99/16180

PRIOR FILING DATE: 1999-07-16

PRIOR APPLICATION NUMBER: 09/350,836

PRIOR FILING DATE: 1999-07-09

PRIOR APPLICATION NUMBER: 09/273,447

PRIOR FILING DATE: 1999-03-19

PRIOR APPLICATION NUMBER: 09/244,444

PRIOR FILING DATE: 1999-02-04

PRIOR APPLICATION NUMBER: 09/122,449

PRIOR FILING DATE: 1998-07-24

PRIOR APPLICATION NUMBER: 09/118,205

PRIOR FILING DATE: 1998-07-16

NUMBER OF SEQ ID NOS: 60

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 42

LENGTH: 14747

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (13641)

OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine

US-09-608-285A-42

Query Match 8.1%; Score 170.8; DB 4; Length 14747;

Best Local Similarity 72.4%; Pred. No. 6.8e-35;

Matches 223; Conservative 26; Mismatches 53; Indels 6; Gaps 3;

QY 1783 GCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAGGATCATTGAGCCCGAG 1842

Db 11074 GCTCAYRCCTRTAATCYAGCACTTTGGGAGGCTGAGGAGGAGGATCATTGAGRTCAG 11015

QY 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAGCCCTGTCTCTACAAAATACAGAGAGAA 1902

Db 11014 RAGTTTTRAGAMYAGCCTGGCCCAACATGGTGAACCCCGTCTCTACTATAAATAACA---AA 10958

QY 1903 AAAATCAGCCAGTCACTGGTGGCATAACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 1962

Db 10957 AAAATTAGCCRGCGGTGGTGGCGCRYGCCCTGTAATCCAGCTACTYGGGAGGCTGAGGCA 10898

QY 1963 GGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCA 2022

Db 10897 GRAGAATCGCTTGAACCCA-GGAGGYGRAGGYTGCACTGAGCYGWGATCRYGCCAYTGCA 10839

QY 2023 CTCCAGCCAGGTGA--CATAGCGAGATCCTGTCTATAAAAAATAAATAAATGAA 2080

Db 10838 CTCCAGCCYRGGMAACAAGAGCRAGACTCCGTCTCAAAAAATAAATAAATAAATAA 10779

QY 2081 CACAGCAA 2088

Db 10778 TAAATAAA 10771

RESULT 9

US-09-557-800C-42/c

; Sequence 42, Application US/09557800C

; Patent No. 6476211

; GENERAL INFORMATION:

; APPLICANT: Ford, John

; APPLICANT: Mulero, Julio

; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE

; POLYPEPTIDES

APPLICANT: Yeung, George

TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 28110/36457

CURRENT APPLICATION NUMBER: US/09/557,800C

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/481,238

PRIOR FILING DATE: 2000-01-11

PRIOR APPLICATION NUMBER: 09/370,265

PRIOR FILING DATE: 1999-08-09

PRIOR APPLICATION NUMBER: PCT/US99/16180

PRIOR FILING DATE: 1999-07-16

PRIOR APPLICATION NUMBER: 09/350,836

PRIOR FILING DATE: 1999-07-09

PRIOR APPLICATION NUMBER: 09/273,447

PRIOR FILING DATE: 1999-03-19

PRIOR APPLICATION NUMBER: 09/122,449

PRIOR FILING DATE: 1998-07-24

PRIOR APPLICATION NUMBER: 09/118,205

PRIOR FILING DATE: 1998-07-16

NUMBER OF SEQ ID NOS: 56

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 42

LENGTH: 14747

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (13641)

OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine

US-09-557-800C-42

Query Match 8.1%; Score 170.8; DB 4; Length 14747;

Best Local Similarity 72.4%; Pred. No. 6.8e-35;

Matches 223; Conservative 26; Mismatches 53; Indels 6; Gaps 3;

QY 1783 GCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAGGATCATTGAGCCCGAG 1842

Db 11074 GCTCAYRCCTRTAATCYAGCACTTTGGGAGGCTGAGGAGGAGGATCATTGAGRTCAG 11015

QY 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAGCCCTGTCTCTACAAAATACAGAGAGAA 1902

Db 11014 RAGTTTTRAGAMYAGCCTGGCCCAACATGGTGAACCCCGTCTCTACTATAAATAACA---AA 10958

QY 1903 AAAATCAGCCAGTCACTGGTGGCATAACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 1962

Db 10957 AAAATTAGCCRGCGGTGGTGGCGCRYGCCCTGTAATCCAGCTACTYGGGAGGCTGAGGCA 10898

QY 1963 GGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCA 2022

Db 10897 GRAGAATCGCTTGAACCCA-GGAGGYGRAGGYTGCACTGAGCYGWGATCRYGCCAYTGCA 10839

QY 2023 CTCCAGCCAGGTGA--CATAGCGAGATCCTGTCTATAAAAAATAAATAAATGAA 2080

Db 10838 CTCCAGCCYRGGMAACAAGAGCRAGACTCCGTCTCAAAAAATAAATAAATAAATAA 10779

QY 2081 CACAGCAA 2088

Db 10778 TAAATAAA 10771

RESULT 10

US-09-608-285A-59/c

; Sequence 59, Application US/09608285A

; Patent No. 6335013

; GENERAL INFORMATION:

; APPLICANT: Ford, John

; APPLICANT: Mulero, Julio

; APPLICANT: Yeung, George

; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE

; POLYPEPTIDES

FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608,285A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583,231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557,800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 59
LENGTH: 15977
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: CD39-L4/L66 Gene Sequence
NAME/KEY: CDS
LOCATION: (245)..(461)
NAME/KEY: CDS
LOCATION: (1454)..(1533)
NAME/KEY: CDS
LOCATION: (2734)..(2877)
NAME/KEY: CDS
LOCATION: (4364)..(4439)
NAME/KEY: CDS
LOCATION: (4679)..(4714)
NAME/KEY: CDS
LOCATION: (5326)..(5414)
NAME/KEY: CDS
LOCATION: (5723)..(5802)
NAME/KEY: CDS
LOCATION: (6751)..(6812)
NAME/KEY: CDS
LOCATION: (7758)..(7859)
NAME/KEY: CDS
LOCATION: (8712)..(8852)
NAME/KEY: CDS
LOCATION: (9831)..(9887)
NAME/KEY: CDS
LOCATION: (11613)..(11728)
NAME/KEY: CDS
LOCATION: (13146)..(13691)
NAME/KEY: CDS
LOCATION: (15702)..(15839)
NAME/KEY: misc feature
LOCATION: (14871)
OTHER INFORMATION: n = a or c or g or t
US-09-608-285A-59

Query Match 8.1%; Score 170.8; DB 4; Length 15977;
Best Local Similarity 72.4%; Pred. No. 7.1e-35;
Matches 223; Conservative 26; Mismatches 53; Indels 6; Gaps 3;
QY 1783 GCTCACACCTGTAATCCTAGCACCTTTGGGAGGCTGAGGAGGAGGATCATTGAGCCCGAG 1842
Db 12304 GCTCAAYRCCTRTAATCYAGCACCTTTGGGAGGCTGAGGAGGAGGATCATTGAGRTCAG 12245
QY 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAACCCCTGTCTCTACAAAATACAGAGAGAA 1902

Db 12244 RAGTTYRAGAMYAGCCTGGCCAACATGGTGAACCCCGTCTCTACTAAATACA---AA 12188
QY 1903 AAAATCAGCCAGTCATGGTGGCATAACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTG 1962
Db 12187 AAAATTAGCCRGCGTGGTGGCGRYGCTGTAAATCCCAGTACTYGGGAGGCTGAGGCA 12128
QY 1963 GGAGGATCACTTGAGCCCGAGGAGTTGGGCTGCAGTGAGCCATGATCACACCACTGCA 2022
Db 12127 GRAGAATCGCTTGAACCCA-GGAGGYGRAGGTGTCAGTGCAGCYGWGATCRYGCCAYTGCA 12069
QY 2023 CTCCAGCCAGGTGA--CATAGCGAGATCCTGTCTAAAAAAATAAAATAATATGGAA 2080
Db 12068 CTCCAGCCYRGMAACAAGAGCRAGACTCCGTCTCAAAAAATAATAATAATAATAA 12009
QY 2081 CACAGCAA 2088
Db 12008 TAAATAAA 12001

RESULT 11
US-09-435-739-42/c
; Sequence 42, Application US/09435739
; Patent No. 6664105
; GENERAL INFORMATION:
; APPLICANT: Pecker, Iris
; APPLICANT: Vlodyavsky, Israel
; APPLICANT: Feinstein, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS
; FILE REFERENCE: 00/20454
; CURRENT APPLICATION NUMBER: US/09/435,739
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 44848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-435-739-42

Query Match 8.0%; Score 170.2; DB 4; Length 44848;
Best Local Similarity 75.4%; Pred. No. 1.8e-34;
Matches 254; Conservative 0; Mismatches 73; Indels 10; Gaps 3;
QY 1783 GCTCACACCTGTAATCCTAGCACCTTTGGGAGGCTGAGGAGGAGGATCATTGAGCCCGAG 1842
Db 23517 GCTCATGCCTGTCTATCCTAGCACCTTTGGGAGGCGGAGGCGGCGGATCATTGAGCCCGAG 23458
QY 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAACCCCTGTCTCTACAAAATACAGAGAGAA 1902
Db 23457 GAGTTCAAGACCAGCCTGGGCAACATGGCAAAACCAATCTCTACAAAAAACA-----CA 23403
QY 1903 AAAATCAGCCAGTCATGGTGGCATAACACCTGTAGTCCCAGCATTC-----CGGGAGGCTGA 1958
Db 23402 AAAATTAGCCGGCATGATGGTAAGCACCTGTAGTCCCAGCTAGCTACTTGGGAGGCTGA 23343
QY 1959 GGTGGGAGGATCACTTGAGCCCGAGGAGGTTGGGCTGCAGTGAGCCATGATCACACCCAC 2018
Db 23342 GGTGGGAGGATCACCTGAGCCCGAGGAGGTCAAGGCTGCAGTGAGCCATGATCATGCCAT 23283
QY 2019 TGCACTCCAGCC-AGGTGACATACGAGATCCTGTCTAAAAAAATAAAATAATAATG 2077
Db 23282 TGCACAGCAGCCTGGGCAACAGATGAGACCCCTGTCTCAAAAAAATAAAATAATAA 23223
QY 2078 GAACACAGCAAGTCCTAGGAAGTAGGTTAAACTAAT 2114
Db 23222 AAAAAACAAAAAAGTAAGAAAAAGAAAAATAATT 23186

RESULT 12
US-09-227-357-100
; Sequence 100, Application US/09227357

